DEPARTMENT OF HEALTH AND HUMAN SERVICES

PUBLIC HEALTH SERVICE

FOOD AND DRUG ADMINISTRATION
CENTER FOR BIOLOGICS EVALUATION AND RESEARCH

Office of Biostatistics and Epidemiology
Division of Biostatistics (HFM-215)

Statistical Review

BLA/Serial Number:	125197 / 0
Drug Name:	Provenge (Sipuleucel-T)
Indication(s):	This autologous active cellular immunotherapy product is indicated for the treatment of men with asymptomatic metastatic androgen independent prostate cancer (AIPC).
Applicant:	*Dendreon Corporation*
Date(s):	11/9/06 (submission) 5/15/07 (BLA first action due)
Review Priority:	Priority
Statistical Reviewer:	*Bo-guang A. Zhen*
Concurring Reviewer:	Ghanshyam Gupta
Medical Reviewer:	Ke Liu
Project Manager:	Lori Tull
cc	Original/HFM-700/Lori Tull DCC/HFM-99 HFM- 215 /Chron HFM- 210 / Mary Foulkes HFM- 219 / Ghanshyam Gupta HFM- 215 / Henry Hsu HFM- 210 / Steven Anderson HFM- 755 / Ke Liu HFM- 755 / Peter Bross

Primary Statistical Reviewer:

Boguang Zhen, Ph.D. — Date: 4/18/07

Concurring Reviewer:

Ghanshyam Gupta, Ph.D. Branch Chief — Date: 4-18-07

Table of Contents

INTRODUCTION

Dendreon is seeking licensure of sipuleucel-T (Provenge®, APC8015) for the treatment of men with asymptomatic, metastatic androgen independent prostate cancer (AIPC). The proposed indication is based upon analyses comparing overall survival between APC8015 treated and placebo groups with the relative absence of significant toxicity in this patient population.

Sipuleucel-T is an autologous active cellular immunotherapy product designed to stimulate an immune response against prostate cancer. APC8015 consists of autologous peripheral blood mononuclear cells (PBMCs), including antigen presenting cells (APCs), that have been activated in vitro with a recombinant fusion protein. The recombinant fusion protein, PA2024, is composed of prostatic acid phosphatase (PAP), an antigen expressed in prostate adenocarcinoma, linked to granulocyte-macrophage colony-stimulating factor (GM-CSF), an immune cell activator.

Sipuleucel-T falls into the class of therapies known as active cellular immunotherapies, sometimes termed therapeutic cancer vaccines. Such immunotherapy products are designed to elicit a specific immune response to a target antigen. While the precise mechanism of action is unknown, sipuleucel-T is designed to induce a cellular immune response targeted against a recombinant fusion protein containing prostatic acid phosphatase (PAP), an antigen expressed in prostate cancer tissue. During ex vivo culture, antigen presenting cells (APCs) take up and process the recombinant target antigen into small peptides that are then displayed on the APC surface. In vivo, T cells bind to and recognize the target antigen peptides on the APC surface, eliciting a response characterized by the proliferation and activation of T cells. These activated T cells are the effector cells thought to be responsible for recognition and destruction of prostate cancer cells in vivo. Sipuleucel-T has been shown to stimulate the proliferation of PAP-specific T cell hybridomas in vitro.

The proposed target indication for sipuleucel-T is for the treatment of men with asymptomatic metastatic androgen independent prostate cancer (AIPC).

All submissions are under **BLA 125197**. The electronic BLA is in CTD format and organized in folders corresponding to BLA structure. Pursuant to the Fast Track Designation agreement and the agreement to submit portions of the application (rolling Biologics License Application), the first portion of the BLA including all clinical and nonclinical sections, draft proposed labeling, and appropriate administrative documents (e.g., forms, table of contents, certifications) was submitted on August 21, 2006. The second/final portion which contains all quality sections along with the final proposed labeling was submitted on November 9, 2006. In addition, Dendreon intends to submit the 4 Month Safety Update in February/March 2007.

STUDY OVERVIEW

Clinical studies of sipuleucel-T have been performed under BB-IND 6933 and include Phase 1, 2, and 3 studies in men with prostate cancer.

Early Phase 1 and 2 clinical studies in men with AIPC were conducted to test the safety and preliminary efficacy of sipuleucel-T (Small 2000, Burch 2000, Burch 2004). The results demonstrated the following: 1) Intravenous infusions of sipuleucel-T in subjects with prostate cancer were generally well tolerated with no dose limiting toxicities observed; 2) Prostate-specific antigen (PSA) reductions of >50% in approximately 10% of subjects were noted, as well as one striking objective response; 3) Three doses of sipuleucel-T resulted in substantial PAP-specific immune responses. Results of open-label Phase 2 trials in men with androgen dependent prostate cancer (ADPC) also demonstrated that intravenous infusions of sipuleucel-T were generally well tolerated with no dose limiting toxicities observed. Additionally, prolongation of PSA doubling time was observed in these studies (Rini 2005a, Rini 2005).

The Phase 3 clinical development program for sipuleucel-T was originally designed based on the results of Phase 1 and 2 trials in men with AIPC and guidance received from the Center for Biologics Evaluation and Research (CBER) and practicing oncologists and urologists. The program consisted of Protocols D9901 and D9902, two identically designed, multicenter, randomized, double blind, placebo-controlled studies in men with asymptomatic, metastatic AIPC. The target number of subjects to be enrolled was 120 in each study, and the primary endpoint for both studies was time to objective disease progression (TTP). Each of the 2 studies was powered to independently meet the primary endpoint of TTP; a pooled analysis of the combined studies was required for sufficient power to meet the secondary endpoint of time to onset of disease-related pain (TDRP). In 1999, no therapy had been shown to prolong survival in men with asymptomatic metastatic AIPC, and there were no Phase 1 or 2 survival data specifically for sipuleucel-T. Therefore, survival was not used as the primary endpoint. However, in both trials, subjects were to be followed until death or until a pre-specified cut-off of 36 months from the time of randomization, whichever occurred first.

The second study (D9902) was initiated shortly after the first study. The trial design, patient eligibility, objectives and statistical considerations were the same as those in the study 1 with a planned sample size of 120 subjects. After the first study results became available showing no overall significance of TTP in 2002, the sponsor did a subset analysis of the study 1 and found that there was a difference of TTP favoring Provenge arm for subjects who had Gleason score $\leq$ 7. At this point, the study 2 had already **enrolled 98 subjects**. The sponsor decided to split the second study into two parts (A and B). Part A (D9902A) contained the initial 98 subjects with Gleason scores ≤ 7 or ≥ 8, and part B to enroll subjects only with Gleason score ≤ 7. The BLA contains data from the two pivotal studies: Studies D9901 and D9902A.

Since data/results from the two pivotal trials were submitted as the main efficacy evidence under this BLA to support the licensing application, the focus of the statistical review is mainly on the two Phase 3 trials (Studies D9901 and D9902A).

STATISTICAL EVALUATION

I. Study D9901

This was a prospective Phase 3, multicenter, double blind, placebo-controlled, randomized trial of immunotherapy with APC8015 for the treatment of subjects with asymptomatic metastatic AIPC. A total of **127** subjects were randomized at multiple investigative centers (**19 clinical study centers**) across the United States. Following a pre-registration and screening process, eligible subjects were randomized to either active treatment or control in a 2:1 ratio (active treatment: control).

An independent third party was employed to generate the randomization schedule for the study. The specifics about the method of randomization employed (block size and degree of imbalance at each study site) were not made known to Dendreon until after the study was unblinded. Subjects were stratified by **clinical study center and use of bisphosphonate therapy** (yes or no) prior to being randomized. The allocation of subjects to treatment utilized multiple blocks, each of size 6, to generate a separate master randomization schedule for each stratum. A centralized, adaptive randomization procedure was employed to maintain the overall enrollment in the study at an approximately 2:1 randomization, while preventing the enrollment at any clinical study center from departing from the 2:1 randomization ratio by a large amount.

Following randomization, subjects from both groups underwent a series of 3 standard leukapheresis procedures (in Weeks 0, 2, and 4), and each procedure was followed 2 days later by infusion of either autologous antigen loaded APCs (APC8015; active treatment) or autologous quiescent APCs without antigen (APC-Placebo; control). The treatment phase of the protocol was complete following the third (Week 4) infusion. For subjects on the control arm, the remaining two-thirds of the quiescent APCs from each leukapheresis that were not used to make APC-Placebo were cryopreserved for possible later use in the preparation of APC8015F for the Phase 2 salvage trial D9903.

At the time that subjects developed disease progression, study treatment could be unblinded. Subjects in the APC8015 group were then treated at the physician's discretion. Subjects in the APC-Placebo group had the option to enter a Phase 2, open label, single-arm salvage trial (Protocol D9903) with a product similar to APC8015. Subjects treated with APC8015 on D9901 were not eligible to participate in the salvage trial.

Regardless of subsequent treatment, subjects without disease-related pain at the time of disease progression were to be followed with weekly pain logs for 4 additional weeks. After subjects developed disease progression, follow-up documentation included treatment-related AEs, first anticancer treatment, and survival for 3 years from the time of randomization or until death, whichever occurred first. Per the statistical analysis plan (Appendix 16.1.9.8), the final analysis of survival was performed 36 months following randomization of the last subject.

1.0 Statistical Analysis Plan

1.1 Efficacy variables

The primary efficacy variable, overall time to disease progression (TTP), was defined as the time from randomization to the first observation of disease progression. Disease progression was defined by any of the following:

Measurable Disease

- A greater than 50% increase in the sum of the products of the perpendicular diameters of all bidimensionally measurable lesions. The change was measured against the smallest sum observed, or compared with baseline if there was no response, using the same techniques as baseline.

- An appearance of at least 2 new lesions or the reappearance of any lesion that had disappeared. All lesions had to have a minimum size of at least 2 cm in 1 dimension to be considered measurable.

Evaluable Disease

- Unidimensionally measurable disease: at least 50% increase in the sum of the measurements of all unidimensionally measurable lesions over the smallest sum observed (over baseline if no response) using the same techniques as baseline.

- Nonmeasurable disease: Clear worsening of nonmeasurable, evaluable disease.

- "Scan only" bone disease: The appearance of at least 2 new areas of abnormal uptake on bone scan. Increased uptake of pre-existing lesions on bone scan did not constitute progression.

Development of Prostate Cancer-Related Events

- The development of a prostate cancer-related event (e.g., spinal cord compression, a pathologic fracture, the development of a requirement for radiation therapy, or other clinically significant disease specific event) constituted progression.

- Failure to return for evaluation due to death or deteriorating condition constituted progression unless the event was clearly unrelated to prostate cancer.

Development of Prostate Cancer-Related Pain

- On the basis of the Investigator's opinion, all of the following criteria had to be met: pain that had the quality and consistency of cancer-related pain, pain that occurred since

enrollment in the trial, and pain that occurred in a location that correlated with a site of cancer, as demonstrated by objective radiographic means.

Secondary efficacy measures included time to development of disease-related pain, objective response rate, and duration of response, time to clinical progression, time to treatment failure, and incidence of Grade 3 or greater treatment-related AEs in all subjects who underwent at least 1 leukapheresis for trial purposes.

Survival as an endpoint was not defined in the protocol and its amendments though the sponsor stated that all subjects would be followed for survival for 36 months after their date of randomization or until death, whichever occurred first. Regarding survival analysis, the sponsor did state that "*This study is not powered to show a survival effect. However, survival data will be summarized descriptively.*" The definition for a survival endpoint was later added to the study report in the BLA submission as:

The survival times for subjects who died during the 3-year follow-up period were defined as the time span (in months) from the date of randomization to the date of death. The survival times for subjects who were alive at the end of the 3-year follow-up period were censored and defined as the time span (in months) from the date of randomization to the censor date of 3 years after the date of randomization.

1.2 Analysis plan

The statistical analysis plan included 2 interim analyses and a final analysis. The first analysis was conducted on data from control subjects only for the purpose of sample size confirmation so no alpha adjustment was made and the sample size was not adjusted based on the results. The second interim analysis was performed on data from 79 subjects from both treatment arms to assess the conditional probability of trial success (i.e., futility analysis) and was conducted at the 0.001 level. The final analysis was conducted at the 0.049 level using the Haybittle-Peto method. Both interim analyses were conducted by an independent third party and Dendreon personnel remained blinded to treatment assignments.

These data were analyzed using the ITT and Safety populations. The ITT population included all randomized subjects, and the Safety population included all subjects who underwent at least 1 leukapheresis. Since all randomized subjects underwent at least 1 leukapheresis in this trial, both populations are identical.

1.3 Analysis of primary efficacy data

The time to disease progression curves were constructed with the Kaplan-Meier technique for the two treatment groups, and the primary null hypothesis (no difference between treatment groups) was tested using the log rank test.

The date of data cut-off for the primary efficacy evaluation was **April 30 2002**. Subjects who did

not experience disease progression by the time of the efficacy analysis were censored at the time of their last known radiographic imaging study. Similarly, if a subject had no disease progression and was lost to follow-up prior to the data analysis, the subject was censored at the date of last radiographic imaging study.

The primary efficacy variable was also summarized by the following subgroups:

- PAP immunohistochemistry expression: 2 subject groupings based on the proportion of cancer cells staining positive for PAP (25% to 74%, = 75%).
- Baseline alkaline phosphatase (within normal limit [WNL] versus above ULN of local reference range).
- Baseline serum PAP levels (WNL of local reference range, > 1 × ULN to < 3 × ULN; = 3 × ULN).
- Prior systemic therapy (castration only, combined androgen blockade, combined androgen blockade plus other; castration plus other was also considered if there were sufficient numbers in the data).

Inferential tests with appropriate adjustment for multiplicity were carried out for the above mentioned variables, but such inferential statistics were only carried out for time to disease progression. Treatment by subgroup interactions were tested to evaluate if any of them represented effect modifiers. These interactions were tested before simple effects and the latter only tested in the event of a significant interaction. Adjustments were made for the 4 interaction tests using the Sidak-Holm method. The tests for simple effects were not considered reportable unless the corresponding interactions were significant (at the $P \leq 0.10$ level).

Subjects were followed for survival for 3 years following their date of randomization. Subjects who were alive at 3 years following randomization were censored at 3 years from their date of randomization. Survival was analyzed using the Kaplan-Meier technique. Survival rate estimates at 3, 6, 9, 12, and every 6 months thereafter and median survival were obtained from the Kaplan-Meier method. Corresponding confidence intervals (CIs) were also computed.

Sample size

The most current version of the protocol (amendment 7, dated 25 JUL 2002) indicates that approximately 120 subjects were planned, with 80 subjects in the APC8015 group and 40 subjects in the APC-Placebo group. A 2:1 randomization was used in order to increase the number of subjects exposed to APC8015. Based on past experience and a review of the literature, a median time to disease progression for subjects treated with APC-Placebo was assumed to be 4 months. A delay in the time to disease progression of 3.7 months (**from 4 to 7.7 months**) was considered clinically significant for subjects with metastatic AIPC. This represents a hazard ratio (HR) of 1.925 assuming an exponential distribution. It was further assumed that accrual into this study would be done within 16 months and that each subject would be followed for up to 3 years.

With a 2-sided 5% level of significance and a 2:1 subject-allocation ratio between the APC8015 and APC-Placebo groups, a total of 80 events was needed to achieve 80% power to detect the specified difference of 3.7 months in median time to disease progression; it was projected that 87 subjects would be sufficient to attain the 80 events (Lachin 1981). To account for non-uniform subject entry and 5% loss to follow-up (Lachin 1986) a total of 96 subjects (64:32 subjects for APC8015:APC-placebo) was necessary.

It was assumed that the same HR would apply to time to disease-related pain. Therefore, a total of 80 pain progression events would also be required to power the disease-related pain endpoint at 80%. However, it was further assumed that 60% of the pain events would be censored when the requisite number of progression events had been recorded. This high level of censoring would result in 32 pain progression events from Study D9901 alone. Increasing the sample size from 96 to 120 would result in a projection of approximately 40 events from pain progression. In order to achieve 80% power and capture 80 pain events for the disease-related pain endpoint, enrollment of 240 subjects was required. Therefore, enrollment of 120 subjects in each of the Phase 3 trials (D9901 and D9902A) was planned with a final pooled analysis for this endpoint.

2.0 Efficacy Evaluation

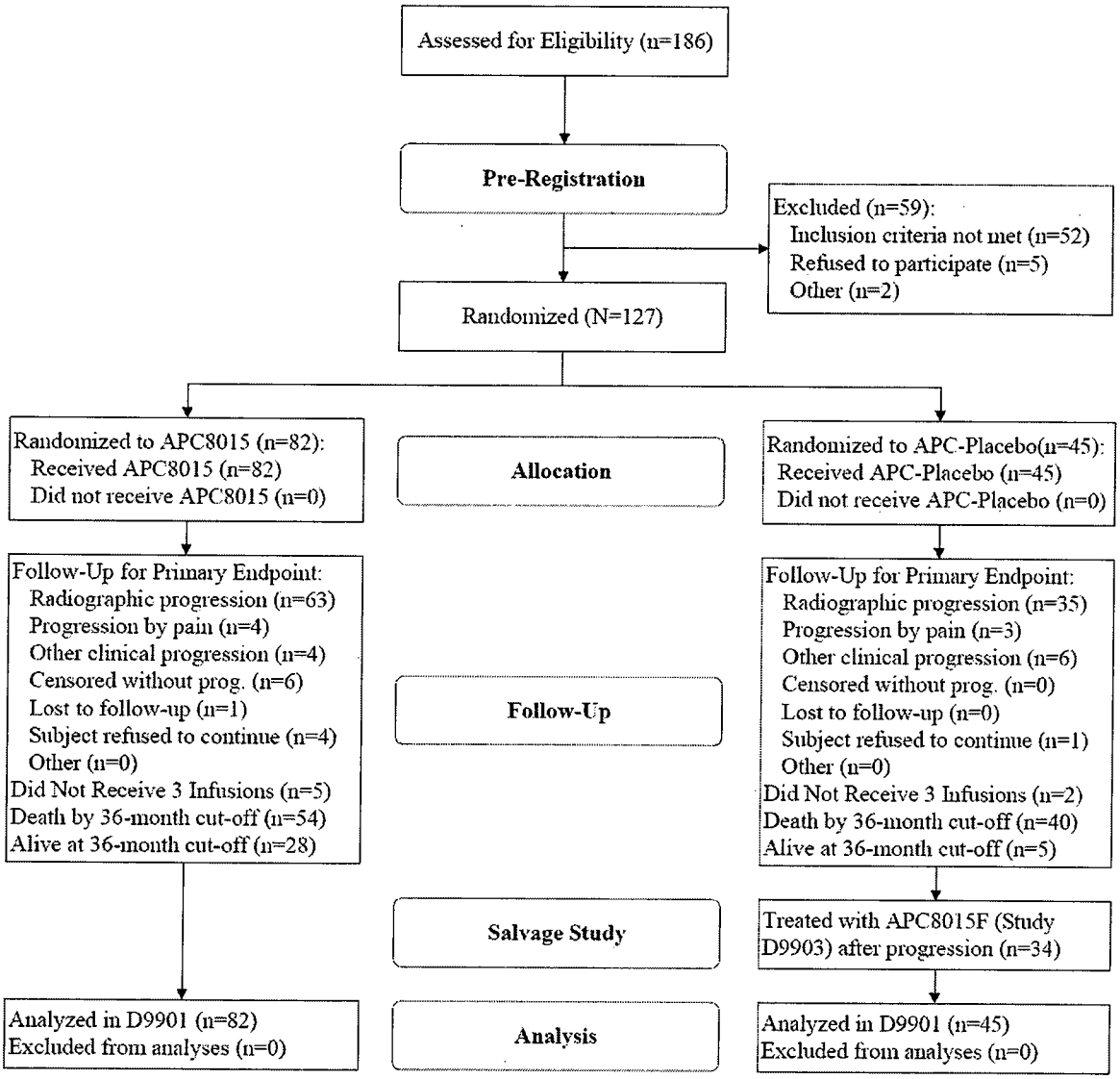
2.1 Disposition of subjects

As shown in Figure 1, 127 of the 186 subjects screened for eligibility were randomized between 04 JAN 2000 and 08 OCT 2001. Of these, 82 subjects were randomized to receive APC8015 and 45 subjects were randomized to receive APC-Placebo. All 127 subjects underwent at least 1 leukapheresis procedure and received at least 1 infusion.

Of the 59 subjects who were screened for the trial but were not randomized, the majority of subjects failed to satisfy the inclusion criteria (52 of 59 subjects, 88%). Five subjects (8.5%) chose not to participate in the trial following their registration visit. Two additional subjects (3.4%) withdrew for other reasons (aortic aneurysm and participation in a separate clinical trial).

Twelve subjects discontinued the 3-year study before completing the trial, but survival at 36 months following randomization was available for all 12 subjects. Four subjects treated with APC8015 and 1 subject treated with APC-Placebo withdrew consent prior to meeting the primary endpoint of disease progression. Rising PSA was not reported as the primary reason for any subject to discontinue the trial.

Figure 1 Schematic of Subject Disposition



Major protocol eligibility deviations occurred for 7.9% of subjects (8 subjects treated with APC8015 and 2 subjects treated with APC-Placebo) and included the following: no evidence of metastatic disease at entry, evidence of pleural effusion at study entry, not medically or surgically castrate at study entry or medical castration therapy discontinued during trial, PSA values demonstrating or confirming androgen independence obtained outside the protocol-specified window, and radiation therapy received during the active period.

2.2 Demographics and other baseline characteristics

A summary of demographics and baseline characteristics is provided in Table 1. The demographic characteristics were similar between the 2 treatment groups. All subjects enrolled in this trial were male, and the majority of subjects were Caucasian (90.6%). The median age in this population was 73.0 years; ages ranged from 47 years to 86 years. The majority of subjects from both treatment groups had a baseline ECOG performance status of 0 (75.6% of subjects treated with APC8015 and 82.2% of subjects treated with APC-Placebo).

Baseline laboratory evaluations were well matched between the treatment groups and are provided in Table 1.

The estimated median time from diagnosis to randomization for subjects treated with APC8015 was 397.6 weeks (approximately 7.6 years) compared to 356.9 weeks (approximately 6.9 years) for subjects treated with APC-Placebo.

Table 1 Summary of Subject Demographics and Baseline Characteristics, ITT

Parameter	APC8015 (N = 82)	APC-Placebo (N = 45)	Total (N = 127[a])	p-value[b]
Age (years)				0.6478
n	82	45	127	
Mean	72.1	71.1	71.7	
SD	8.1	8.3	8.1	
Median	73.0	71.0	73.0	
Minimum	47.0	50.0	47.0	
Maximum	85.0	86.0	86.0	
Race, n (%)				0.1405
Caucasian	73 (89.0)	42 (93.3)	115 (90.6)	
African American	8 (9.8)	1 (2.2)	9 (7.1)	
Hispanic	1 (1.2)	1 (2.2)	2 (1.6)	
Unknown	0 (0.0)	1 (2.2)	1 (0.8)	
Weight (lbs)				0.1615
n	82	44	126	
Mean	199.9	191.2	196.9	
SD	33.3	31.5	32.8	
Median	194.1	186.5	192.0	
Minimum	141.5	135.0	135.0	
Maximum	334.4	272.1	334.4	
ECOG Performance Status, n (%)				0.2953
0	62 (75.6)	37 (82.2)	99 (78.0)	
1	20 (24.4)	8 (17.8)	28 (22.0)	
Serum PSA (ng/mL)				0.9725
n	81	45	126	
Mean	181.8	168.0	176.9	
SD	469.6	422.9	451.8	
Median	46.0	47.9	47.3	
Minimum	3.5	7.9	3.5	

Parameter	APC8015 (N = 82)	APC-Placebo (N = 45)	Total (N = 127[a])	p-value[b]
Maximum	3621.0	2799.0	3621.0	
Serum PAP (ng/mL)				0.9908
n	74	40	114	
Mean	17.4	17.6	17.5	
SD	32.8	30.0	31.7	
Median	7.0	6.5	6.8	
Minimum	0.7	0.3	0.3	
Maximum	250.5	163.0	250.5	
Alkaline Phosphatase (U/L)				0.5523
n	82	45	127	
Mean	151.3	131.2	144.2	
SD	160.5	118.5	146.8	
Median	102.0	92.0	96.0	
Minimum	42.0	38.0	38.0	
Maximum	1233.0	627.0	1233.0	
LDH (U/L)				0.8146
n	80	43	123	
Mean	184.4	186.9	185.3	
SD	54.8	64.7	58.2	
Median	173.5	172.0	172.0	
Minimum	119.0	108.0	108.0	
Maximum	533.0	453.0	533.0	
Hemoglobin (g/dL)				0.5516
n	82	45	127	
Mean	13.0	12.8	12.9	
SD	1.3	1.3	1.3	
Median	13.0	13.1	13.1	
Minimum	8.5	9.3	8.5	

Parameter	APC8015 (N = 82)	APC-Placebo (N = 45)	Total (N = 127[a])	p-value[b]
Maximum	16.5	14.8	16.5	
Time from Diagnosis to Randomization (weeks)				0.5947
n	82	45	127	
Median	397.6	356.9	370.4	
Minimum	40	84	40	
Maximum	903	973	973	

[a] Unknown values were not categorized; therefore, all parameters do not represent 127 subjects.
[b] p-value compares treatment groups using a Cochran-Mantel-Haenszel test.

The protocol required subjects to have a tumor specimen (tissue block, core biopsy, or pre-cut unstained slides) submitted to a central pathology facility for immunohistochemistry testing of PAP. Eligibility required a positive PAP immunohistochemistry reaction in ≥ 25% of cells. The PAP immunohistochemistry results are summarized in Table 2. A higher percentage of subjects in the APC8015 group than in the APC-Placebo group had tumor specimens with at least 75% PAP-positive cells.

Table 2 Summary of PAP Immunohistochemistry, ITT

	APC8015 (n = 82)	APC-Placebo (n = 45)	Total (N = 127)	p-value[a]
Percent PAP Positive, n (%)				0.0618
25% - 74%	19 (23.2)	17 (37.8)	36 (28.3)	
≥ 75%	63 (76.8)	28 (62.2)	91 (71.7)	

[a] p-value compares treatment groups using a Cochran-Mantel-Haenszel test.

For this study, 116 of 127 subjects (91.3%) had a Gleason score assigned by a central pathology laboratory prior to randomization. The Gleason scores obtained by local pathology facilities were used for those subjects not given Gleason scores by the central pathology facility. Gleason score and tumor status at baseline is presented in Table 3. Overall, a majority of the subjects had a Gleason score ≤ 7 (75 subjects [59.1%] versus a Gleason score ≥ 8 (52 subjects [40.9%]). Differences between the treatment groups were not statistically significant.

Table 3 Summary of Gleason Score and Tumor Status at Baseline, ITT

	APC8015	APC-Placebo	Total	p-value[a]
Gleason Score				
N	82	45	127	
Median	7	7	7	
Minimum	4	5	4	
Maximum	10	10	10	
Score, n (%)				0.1875
≤ 6	22 (26.8)	7 (15.6)	29 (22.8)	
= 7	28 (34.1)	18 (40.0)	46 (36.2)	
≥ 8	32 (39.0)	20 (44.4)	52 (40.9)	
Localization of Disease[b]				0.0538
N	81	42	123	
Bone metastases only	34 (42.0)	10 (23.8)	44 (35.8)	
Soft tissue metastasis/pelvis recurrence only	5 (6.2)	3 (7.1)	8 (6.5)	
Both bone metastases and soft tissue metastasis/pelvic recurrence	42 (51.9)	29 (69.0)	71 (57.7)	
Number of Bone Metastases per subject				0.3447
N	82	45	127	
0	5 (6.1)	4 (8.9)	9 (7.1)	
1 – 5	31 (37.8)	17 (37.8)	48 (37.8)	
6 – 10	12 (14.6)	12 (26.7)	24 (18.9)	
> 10	34 (41.5)	12 (26.7)	46 (36.2)	

[a] p-value compares treatment groups using a Cochran-Mantel-Haenszel test controlled by center for categorical data and a 2-way analysis of variance with factors of treatment group and center for continuous data.

[b] Localization of disease could not be determined for 4 subjects due to lack of baseline scans for soft tissue disease.

It should be noted that p-values provided by the sponsor in the above tables should not be considered as those from a hypothesis test for the difference between the two arms. They just reflect the chance of obtaining the observed difference (and more extreme) between the two arms in these demographic and baseline characteristics factors when in fact the two samples were randomly drawn from the same population.

2.3 Efficacy results

The protocol prospectively designated time to disease progression as the primary endpoint and specified that complete survival data (up to 36 months) would be collected. Survival is the focus of this study report. Survival is objectively ascertained, represents the standard for establishing clinical benefit in oncology clinical trials, and best represents the therapeutic effect of APC8015. To be consistent with the prospectively defined protocol, time to disease progression and the secondary endpoints are presented first in this study report, but the most extensive information and critical analyses are focused on survival.

The efficacy review focuses on the survival endpoint.

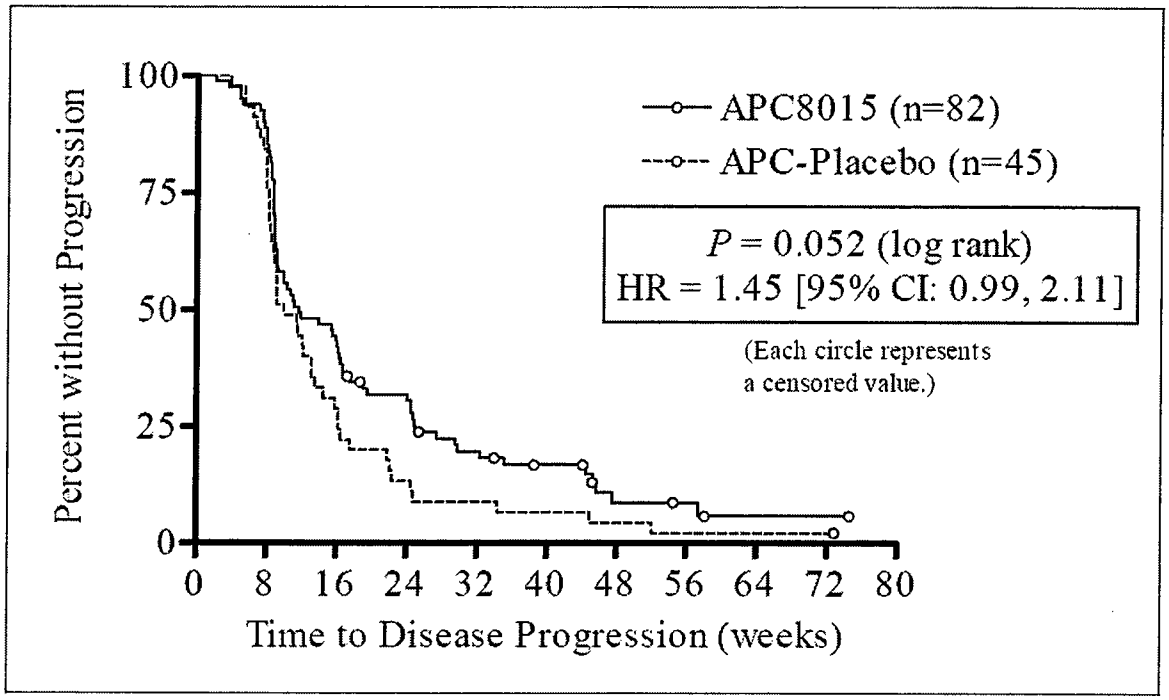
2.3.1 The primary endpoint

The sponsor states in this BLA that the primary efficacy endpoint was the overall time to disease progression. The following analyses for the primary endpoint presented in the study report by the sponsor were based on the unblinded review data. The terminology change for the primary endpoint and the adequacy of using unblinded review data will be discussed later in the section "Statistical Findings and Comments".

Out of 127 subjects randomized on this study, 115 subjects (90.6%) contributed a progression event. Ninety-eight subjects (77.2%) had progression documented by imaging (as determined by independent, blinded, radiology review). Of the 98 subjects who progressed based on imaging studies, 48 subjects treated with APC8015 and 24 subjects treated with APC-Placebo progressed based on bone disease, while 15 subjects treated with APC8015 and 11 subjects treated with APC-Placebo progressed based on soft tissue disease. Ten subjects (7.9%) had progression based on clinical events other than radiographic events, as defined in the protocol, and 7 subjects (5.5%) had progression based on the onset of cancer-related pain. Additionally, 6 subjects (4.7%) were censored prior to meeting the disease progression endpoint and 6 subjects (4.7%; all in the APC8015 group) were censored without disease progression at the primary efficacy evaluation cut-off date (30 APR 2002).

When the Kaplan-Meier curves for time to disease progression were compared, there was a delay from randomization to disease progression in the APC8015 group compared with the APC-Placebo group (P = 0.052, log rank; unadjusted HR = 1.45 [95% CI: 0.99, 2.11]; Figure 2). After Week 8 the Kaplan-Meier curves showed a marked separation that persisted throughout the remainder of follow-up. The estimated median time to disease progression was 11.7 weeks in the APC8015 group compared with 10.0 weeks in the APC-Placebo group.

Figure 2 Primary Efficacy Endpoint, Time to Disease Progression (Kaplan-Meier Method), ITT



The sponsor also reported time to objective disease progression confirmed by imaging studies as follow (p74 of the Study Reports for D9901): During the trial, available imaging studies were evaluated for all subjects by a central, independent radiology facility. A supplementary analysis was conducted on the time from randomization to objective disease progression confirmed by imaging studies. Results of this analysis indicated that statistical significance was not reached in the ITT population (P = 0.183, log rank; unadjusted HR = 1.32 [95% CI: 0.87, 2.00]).

2.3.2 The secondary endpoints

There was no statistically significant difference between the treatment groups with respect to time to onset of disease-related pain progression (P = 0.210 log rank; unadjusted HR = 1.47 [95% CI: 0.80, 2.68]). The median time to onset of disease-related pain in subjects treated with APC-Placebo was estimated to be 24.0 weeks, while the median time to onset of disease-related pain in subjects treated with APC8015 was not estimable. The pain-free rate estimate at 12 weeks, 71.5% of the subjects treated with APC8015 and 69.7% of the subjects treated with APC-Placebo had not experienced onset of disease-related pain progression.

Since no subjects experienced a tumor response based on review by the central radiology facility, there is no evaluation on tumor response rate and duration of response.

Time to clinical progression was analyzed to determine the difference in the primary endpoint in cases where both subjective evidence and independently confirmable evidence of disease

progression were present. For the time to clinical progression analysis, the first evidence of disease progression for each subject was used, whether based on subjective or independently confirmable evidence. (For the primary endpoint of time to disease progression, the date of independently confirmable disease progression was used when available.) Twenty-two subjects treated with APC8015 and 18 subjects treated with APC-Placebo had a clinical progression date that differed from their time to disease progression date. There was a trend toward a prolonged time from randomization to clinical progression in the APC8015 group compared with the APC-Placebo group, which approached but did not reach statistical significance (P = 0.061, log rank; unadjusted HR = 1.44 [95% CI: 0.98, 2.10]).

Time to treatment failure was defined as the time from randomization until any of the following occurred: disease progression, death, or withdrawal for any reason except withdrawal of consent. (Withdrawal of consent caused the subject to be censored at the time of the last visit.) Initiation of other primary anticancer therapy, including radiation therapy, in the absence of study withdrawal was considered treatment failure for the purpose of this endpoint, as of the date the therapy was initiated. The difference between APC8015 and APC-Placebo in time to treatment failure was not statistically significant (P = 0.124, log rank; unadjusted HR = 1.34 [95% CI: 0.92, 1.94]).

2.3.3 Overall survival

The primary comparison between two arms in overall survival WAS NOT pre-specified in the protocol and the statistical analysis plan before unblinding the data.

The analysis of the 3-year survival data was based on the ITT population of all 127 randomized subjects. Every subject was followed until death or the pre-specified cut-off of 36 months(i.e.: the cutoff date should be in October 2004 since the last patient was enrolled in October 2001); there were no censored events prior to the 36th month of follow-up. Although there was one subject who lost to follow-up as shown in Figure 1, this is just specific for the time to disease progression endpoint.

The first analysis of the survival data were based on the Kaplan-Meier technique and the log rank test. Subjects treated with APC8015 demonstrated an improvement in overall survival, compared to those treated with APC-Placebo (P = 0.010, log rank; Figure 3). The unadjusted HR was 1.71 (95% CI: 1.13, 2.58), indicating a 41% reduction in the death rate for subjects treated with APC8015 compared to APC-Placebo. The median survival time for subjects treated with APC8015 was 4.5 months longer than that for subjects treated with APC-Placebo (median survival times of 25.9 months and 21.4 months, respectively). Table 4 and Table 5 present key summary statistics that characterize the differences between the 2 treatment arms.

Figure 3 **Overall Survival (Kaplan-Meier Method), ITT**

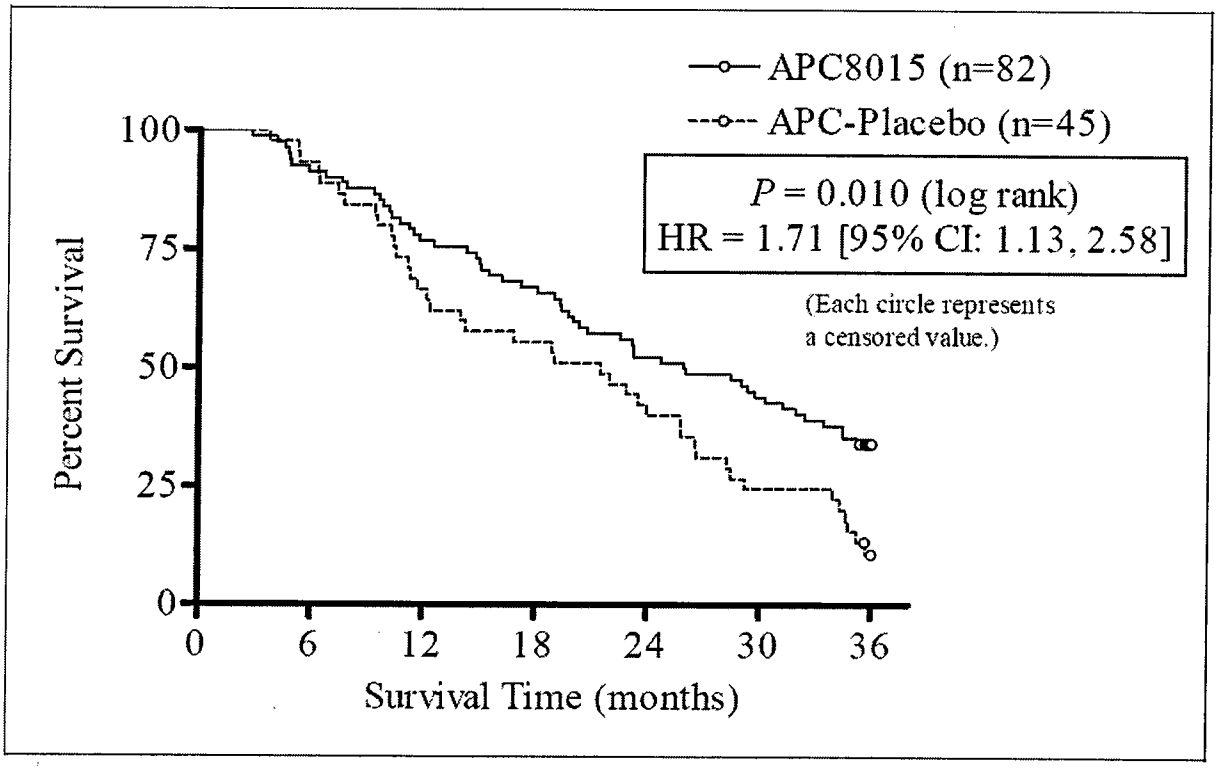


Table 4 **Summary Statistics, ITT**

Treatment	Number of Subjects	Deaths Over 36-Month Follow-Up	Subjects Alive at 36-Month Visit	Death Percentiles, months		
				25%	50%	75%
APC8015	82	54	28	14.3	25.9	≥ 36.0
APC-Placebo	45	40	5	10.5	21.4	29.2

Table 5 **Kaplan-Meier Survival Rate Estimates, Percent ITT**

Treatment	6 months	12 months	18 months	24 months	30 months	36 months
APC8015	91.5	76.8	67.1	52.4	43.9	34.1
APC-Placebo	93.3	66.7	55.6	40.0	24.4	10.7

Following study closure, Dendreon attempted to obtain death certificates and other source documents to confirm the cause of death. Based on a review of these additional documents as well as data obtained from the death summary CRF, all causes of death were ascertained. Given the importance of the death date, Dendreon also compared the death date recorded by the clinical study center (on the Death Summary CRF) to the date listed on the Social Security Death Index

(SSDI) for 93 of the 94 subjects who died during the 36 month follow-up (An SSDI death date was not available for 1 subject). In the majority of these cases (86 of 93 cases) the death dates from the 2 sources were identical. Discrepancies were noted for 4 subjects treated with APC8015 and for 3 subjects treated with APC-Placebo. The observed differences are minor and would not substantively change the survival difference observed between APC8015 and APC-Placebo (Detailed results of these findings are contained in Section 12.3.3.1).

Several other sensitivity analyses were performed to test the robustness of the survival results. Specifically, these sensitivity analyses included the following:

- Removal of influential subjects
- Removal of investigational study centers
- Reversing the treatment assignment of subjects with randomization errors
- Removal of subjects with protocol deviations
- Assessing the influence of cell processing centers (CPCs) on survival
- Comparison of use of chemotherapy during long-term follow-up
- Assessing the influence of prognostic factors on the observed survival effect

Six of the longest surviving subjects in the APC8015 group had to be removed before the p-value exceeded 0.05.

The exclusion of study center 69 was the only one that resulted in a non-significant p-value (P = 0.062, log rank). This was the largest study center with a total of 20 subjects (15.8% of all subjects). Each of the other study center exclusions yielded a significant survival finding.

2.3.4 Proportional model for survival analysis

The sponsor evaluated the individual effect of the 21 potential prognostic factors as listed in Appendix 16.1.9.11 and found that eight of these prognostic factors (age, alkaline phosphatase, hemoglobin, lactate dehydrogenase [LDH], localization of disease, number of bone metastases, PSA, and weight) could be independently identified as predictors of survival at the 0.05 level. The sponsor also stated that each one of these 8 variables has previously been identified in the literature as a significant prognostic factor of survival. Additional prognostic factor, serum PAP, was also identified as a significant prognostic factor by the sponsor, but the sponsor excluded it from the model since they believed that it had substantially more missing data.

In order to build a model that was predictive of survival, the 9 prognostic factors (including serum PAP) identified in the univariate analyses were considered as candidates and included in a multivariate PHR model.

The backwards stepwise selection method (P = 0.05 for entry and P = 0.10 for removal, likelihood ratio test) was then used to identify the prognostic factors that added significantly to the fit of the model. Serum PAP was found not to be significant following the backwards elimination procedure. This analysis was repeated without serum PAP as a covariate since there was a relatively large number of missing PAP values. The results of this analysis reduced the

number of prognostic factors remaining in the model to 5. The 5 baseline prognostic factors that remained in the final model were LDH (ln), PSA (ln), localization of disease, number of bone metastases, and body weight (lbs). It should further be noted that the treatment effect continued to be significant at every step of the backward elimination procedure and was a predictor of survival in the final model. Following identification of these 5 prognostic factors and in order to utilize all of the data available, the PHR analyses were conducted with just these 5 variables and the treatment effect in the model because the 3 eliminated variables had missing data. The results of these analyses are presented in Table 6. Note that the p-value for localization was greater than 0.05 but less than 0.10. The treatment effect was significant (P = 0.002, Wald.s test; adjusted HR = 2.16).

Table 6 Proportional Hazards Regression Model of Survival – Cox model (I)

		95% CI for HR		
	HR	**Lower**	**Upper**	**p-value**
Treatment with APC8015	2.16	1.33	3.50	0.002
Baseline PSA (ln)	1.32	1.09	1.59	0.004
Baseline LDH (ln)	4.89	2.01	11.88	< 0.001
Localization of Disease	1.54	0.96	2.46	0.072
Number of Bone Metastases				0.005
Lesion count: 0-5 lesions versus 6-10 lesions	1.74	0.93	3.25	
Lesion count: 0-5 lesions versus > 10 lesions	2.30	1.37	3.85	
Weight	0.99	0.98	1.00	0.032

N = 127: Events = 85: Censored = 32, and cases with missing values = 10
Abbreviations: CI = confidence interval; HR = hazard ratio; ln = natural log

2.3.5 Cell dose and product potency

While the design of the study was not one in which the cell dose was specified, the analysis of cell dose and product potency was requested as a means of examining the relationship between survival and cell dose. To this end, survival data for subjects treated with APC8015 were assessed in the context of the key release specification parameters of the product, notably, the TNC, CD54 cell count, and the upregulation of CD54. CD54 cell count and CD54 upregulation were chosen as biologically relevant release specifications because of CD54's uniform expression on APCs, its role in the immunologic synapse between APCs and T cells, and its role as a marker of APC activation. Specifically, experiments have demonstrated that the CD54+ population of mononuclear cells possesses the ability to take up the PA2024 antigen and present epitopes of PAP to hybridoma cell lines recognizing PAP epitopes.

In the simple Cox PHR model, subjects who had a CD54 cell count at or above the median of 2.5 x 10^9 cells, or CD54 upregulation ratio at or above the median of 23.3, had an improved survival compared to those subjects below the median (HR = 0.63 and HR = 0.79, respectively). A

significant effect on survival was observed between subjects above and below the median TNC count of 10.8 x 10^9 cells (HR = 0.52; P = 0.018; Table 7). A multivariate Cox PHR model was used to determine whether cell counts correlated with survival when correcting for the 5 key prognostic variables of baseline PSA, lesion count, localization of disease, baseline LDH, and weight. A similar trend was observed when TNC was included in the multivariate Cox PHR model (P = 0.054; HR = 0.56). As in the simple Cox model, there was no statistically significant correlation for total CD54 cell count (P = 0.233; HR = 0.70) or CD54 upregulation (P = 0.274, HR = 0.72) in the multivariate Cox PHR model.

It should be noted that there was no information on the cell dose and characteristics from the placebo group and the study was not designed to provide confirmative evidence for relationship between survival and cell dose. The significant result below may just indicate that TNC count is another patient factor predicting prognosis.

Table 7 Analyses of Survival and Cell Dose

Variable	N	HR	p-value
CD54 cell count (< 2.5 x 10^9 cells versus ≥ 2.5 x 10^9 cells)	82	0.63	0.098
CD54 upregulation (< 23.3 versus ≥ 23.3)	82	0.79	0.395
TNC count (< 10.8 x 10^9 cells versus ≥ 10.8 x 10^9 cells)	82	0.52	0.018

Abbreviations: HR = hazard ratio; TNC = total nucleated cell

2.3.6 Comparison of Use of Chemotherapy during Long-Term Follow-Up

In order to assess whether the use of chemotherapy, in particular the use of docetaxel, influenced the survival results, the two treatment groups were compared with respect to the administration of chemotherapy use following protocol treatment. This analysis revealed that there was not a statistically significant imbalance in the administration of chemotherapies between treatment groups. Based on data from 122 of 127 randomized subjects, **54.4%** of subjects who received APC8015 and **62.8%** of subjects who received APC-Placebo were treated with any chemotherapy following disease progression (Table 8). (Only the type of chemotherapy and the date of initiation were collected for this trial. No data on the dose or duration of chemotherapy are available.)

Table 8 Chemotherapy Use Following Therapy

Chemotherapy	APC8015 (n = 78)	APC-Placebo (n = 41)	p-value (Fisher's Exact)
Docetaxel	29 (37.2%)	20 (48.8%)	0.244
Chemotherapy other than taxanes	34 (43.6%)	13 (31.7%)	0.240
Taxane-based chemotherapy	34 (43.6%)	22 (53.7%)	0.337
Any chemotherapy[a]	43 (54.4%)	27 (62.8%)	0.445

[a] For any chemotherapy, APC8015 (n = 79) and APC-Placebo (n = 43)

3.0 Safety Evaluation

All 127 subjects randomized on this trial underwent at least 1 leukapheresis procedure and are included in the safety population. A total of 120 of 127 treated subjects (94.5%) received all 3 infusions. Five subjects (3.9%), 3 subjects treated with APC8015 and 2 subjects treated with APC-Placebo, received 2 infusions. Two subjects (1.6%), both of whom were treated with APC8015, received only 1 infusion during the study. Only 1 subject did not receive all infusions as a result of treatment-related AEs (hypoxia, fever, and chills). Two subjects (9126-154 treated with APC8015 and 9126-090 treated with APC-Placebo) did not receive all 3 infusions following leukapheresis-related sepsis.

The majority of subjects in this study experienced mild to moderate AEs. Events that occurred most frequently were generally mild and resolved within 1 to 2 days. Grade 3 or 4 AEs were observed in a total of 26 subjects (31.7%) in the APC8015 treatment arm and 12 subjects (26.6%) in the APC-Placebo treatment arm. Only 5 subjects experienced SAEs that were considered related to study treatment, 2 subjects treated with APC8015 and 3 subjects treated with APC-Placebo.

A total of 94 deaths occurred within the 3-year reporting period, with a lower percentage of deaths in the APC8015 group versus the APC-Placebo group (66% deaths versus 89%, respectively). No deaths occurred within 30 days of the last product infusion.

Detailed review on safety can be seen in the medical reviewer's memo.

4.0 Statistical Findings and Comments

4.1 Overall Survival Endpoint

The statistical review focuses mainly on overall survival since the key efficacy evidence is based on findings of the difference between APC8015 treated patients and placebo patients in overall survival.

4.1.1 Difficulty in interpreting the hypothesis test results

Study D9901 was unblinded in June 2002. In the study protocol and its amendments including Amendment #6 (9/27/2001) and Amendment #7 (7/25/2002, last amendment), overall survival as an endpoint was NOT defined. The definition for a survival endpoint was later added to the study report under this BLA submission as:

The survival times for subjects who died during the 3-year follow-up period were defined as the time span (in months) from the date of randomization to the date of death. The survival times for subjects who were alive at the end of the 3-year follow-up period were censored and defined as the time span (in months) from the date of randomization to the censor date of 3 years after the date of randomization.

However, the sponsor cited the following paragraph from the protocol and made a statement in the proposed labeling: "*Overall survival at 36 months was a pre-specified analysis*". The sponsor believes that survival analysis (as bolded below) was mentioned in Section 9.6.2.1 so the results presented in labeling could be considered as from a pre-specified analysis.

9.6.2.1 Primary endpoint

The analysis of time to disease progression will be conducted on the ITT and Efficacy Evaluable Populations. The primary analysis will be on the ITT population.

The primary null hypothesis is that the time to disease progression curve of the APC8015 group is not different from that of the control group. The corresponding alternative hypothesis is that there is a difference. Time to disease progression curves will be constructed using Kaplan-Meier technique (25) for the two treatment groups and the primary hypothesis tested using the log-rank test.

As supporting analyses, estimates of ***survival rate*** *and progression free frequencies at six, nine, twelve and eighteen, twenty four and thirty six months, and* ***median survival*** *will be provided based on the Kaplan-Meier curves, with corresponding confidence intervals; and the Cox proportional hazards model may be used to adjust for prognostic variables (26).*

A secondary analysis using a stratified log-rank test or Cox regression may be applied to adjust for treatment by center interaction if indicated. In addition, treatment interaction with centers

pooled by site of cell-processing will be explored, as indicated.

In addition, the type-specific hazards for the pre-specified types of disease progression may be estimated via Cox models if there is strong interest suggested by the data.

It appears to this reviewer that Section 9.1.2.1 was describing the primary and supportive analyses for the primary endpoint, time to disease progression, since no survival endpoint was defined in the protocol and its amendments. It is unclear whether the so-called "*survival rate*" and "*median survival*" just refer to the survival analysis methods for analyzing the primary endpoint or have something to do with a survival endpoint. If this is for a survival endpoint, what survival endpoint the sponsor intended to analyze remains everybody's guess (it could be an overall survival, a cause-specific survival, a treatment-relative survival endpoint, etc.). Even if overall survival, i.e.: all deaths were counted as an event, had been considered as the survival endpoint, it would still have several ways to define the endpoint – time from diagnosis to death, time from randomization to death, time from first treatment to death, etc.

Even if (this is the second even if) one had clearly defined the overall survival as time from the date of randomization to the date of death as the sponsor did later in their submission, the primary analysis method on this endpoint would still need to be pre-specified (use of log-rank test with or without a stratification factor? Which set of covariates would be used in Cox model?). Therefore, the analysis on overall survival at 36 months should not be interpreted as 'pre-specified analysis'.

Nevertheless, no matter how Section 9.6.2.1 can be interpreted regarding survival analysis, it is very clear from Section 9.3 in the protocol below that the sponsor did not plan any hypothesis test for an effect for overall survival or any other survival endpoints. This indicates that the sponsor did not pre-specify a statistical analysis method for the primary comparison between the two arms in overall survival—the comparison resulted in key efficacy evidence in support of the licensing application.

9.3 Analysis endpoints

.....

This study is not powered to show a survival effect. However, survival data will be summarized descriptively.

Thus, one would face the challenge of estimating type I error rate (the probability of making a false positive claim for treatment effect) and statistical correction for multiplicity (multiple endpoints and multiple analysis methods) due to the un-prespecified nature in survival analysis. The following are the summaries for survival analysis:

a. Overall survival as an endpoint was not defined in the protocols;
b. A statistical analysis method for the primary comparison of overall survival between the two arms was not pre-specified;
c. The post-hoc analysis makes it very difficult to interpret hypothesis test results for overall survival.

4.1.2 Additional sensitivity analyses using Cox model

The log-rank test for treatment effect resulted in a p-value of 0.01 and the p-value was reduced to 0.002 when adjusting for a set of covariates using Cox proportional hazards regression (PHR) model as reported by the sponsor [Cox model (I) in Table 9]. The results were verified by the statistical reviewer.

It should be noted that there are so many ways to incorporate covariates in Cox PHR model. These ways include, but are not limited to: 1) criteria of selecting candidate covariates; 2) covariate selection procedure in the model (forward, backward, stepwise, etc.); 3) type of scale used for non-binary candidate covariates – using original scale, using log scale, using median as cut-off point, using other cut-off point, etc. Therefore, it is very difficult to judge which model should be treated as the primary when an analysis was not pre-specified because anyone would tend to select the one for his/her own favor intentionally or unintentionally.

Several additional sensitivity analyses with different set of covariates in Cox model were conducted by the statistical reviewer and the results are displayed in Table 9. In one of the analyses [Cox model (II)] when using localization of disease (bone and soft tissue vs. bone only or soft tissue only), PSA (<20, 20-<100, ≥100), and Gleason score (≤6, 7, ≥8) as covariates, p-value=0.0778 (confirmed by the sponsor). As shown in this table, analysis using different set of covariates resulted in different p-value.

Table 9 Treatment effect on overall survival using Cox model with different set of covariates (total subjects: 127)

Testing Method	Hazard Ratio	P-value	Covariate Adjustment	#Patients Missing
Log-rank test*	1.71	0.01	None	0
Cox Model (I)*	2.16	0.002	Localization of disease, PSA(ln), LDH(ln), Weight, #Bone metastases (≤5, 6-10, >10)	10
Cox Model (II)~	1.47	0.078	Localization of disease, Gleason Score (≤6, 7, ≥8), PSA (<20, 20 - <100, ≥100)	5
Cox Model (III)~	1.51	0.060	Localization of disease, PSA (<20, 20 - <100, ≥100)	5
Cox Model (IV)~	1.53	0.053	Localization of disease, Gleason Score (≤6, 7, ≥8)	4
Cox Model (V)~	1.52	0.048	Gleason Score (≤6, 7, ≥8), PSA (<20, 20 - <100, ≥100)	1

*: The sponsor's analyses and confirmed by the reviewer
~: Reviewer's analysis

Although 36-month survival data are complete for all randomized subjects, there are some subjects with missing covariate data mainly due to the nature of un-prespecification in overall survival analysis. It should be noted that 10 subjects were excluded from the analysis due to missing covariate data when Cox model (I) was used by the sponsor (resulting in a reduction of

p-value from 0.01 to 0.002). Therefore, the statistical reviewer further analyzed the survival data by breaking subjects into two groups: 1) subjects with any missing covariate data and were excluded from the analysis in Cox model (I) by the sponsor; 2) subjects with complete covariates data for Cox model (I). As presented in Table 10, one can see that patients without complete covariate data had poor median survival in APC8015 treated group, compared to the rest of the APC0805 treated patients. In contrast, placebo patients with missing covariate data had longer median survival relative to the rest of the placebo patients. The randomization principle may be seriously compromised when using Cox model (I) and exclusion of subjects can lead to biased estimates of treatment effects. Thus, the results from any analyses with exclusion of subjects should be interpreted cautiously.

Table 10 Impact of missing covariate data in Cox Model (I)

Complete Covariate Data	APC8015			Placebo		
	N	# Deaths	Median Survival (mons)	N	# Deaths	Median Survival (mons)
No	4	4	19.7	6	5	22.1
Yes	78	50	28.7	39	35	19.0

In order to include all randomized patients in the analysis using Cox Model (I), one may want to impute missing covariate data using different imputation approaches. It is certain that different approaches will result in different answers. For example, one may want to assign a median value estimated from those who have PSA or LDH data to those with missing PSA or LDH values, respectively. If further assume that the 4 patients without "Localization of disease" information had bone and soft tissue disease, the hazards ratio for treatment effect was 1.52 (p=0.054) using Cox Model (I). If assume that the 4 patients without Localization of disease information had bone only or soft tissue only disease, the hazards ratio for treatment effect was 1.57 (p=0.036). These two imputation approaches did not low down the p-value compared to the log-rank test (no covariate adjustment). Again, since no imputation approaches was pre-specified, anyone would tend to select her/his favorite one as the primary intentional or unintentionally.

The reviewer also evaluated all demographics and baseline characteristics parameters presented in the study report for their impacts on overall survival. Column "**Cov. Effect only**" in Table 11 shows the individual effect of each covariate on survival. From this column, one can see that age, weight, baseline PSA, baseline PAP, baseline hemoglobin, localization of disease, baseline LDH, and number of bone metastases had statistically significant (without adjustment for multiplicity) impacts on overall survival. The third column in this table shows the treatment effect while adjusting each of the covariates. P-values for treatment effect ranged from 0.003 to 0.041 depending on which covariate was adjusted in the model.

It should also be noted that the proportional hazards assumption as required for Cox PHR model may be violated. From Figure 4 below, one can see that there was a minimal difference in the two curves within the first ten months and the hazards between the two arms may not be proportional over time. This may indicate that Cox PHR model may not be the optimal analysis

tool for this type of survival data.

Table 11 Single covariate effect and treatment effect on overall survival in Cox model

Selection of Covariate	Cov. Effect only HR	Cov. Effect only P-value	Treatment Effect HR	Treatment Effect P-value	#Patients Missing
Baseline PSA (ln)	1.36	<.0001	1.67	0.015	1
Baseline PSA		0.0006	1.59	0.028	1
<20 vs. 20 - <100	1.58				
<20 vs.≥100	3.06				
#Bone metastases		0.0017	1.89	**0.003**	0
0-5 vs. 6-10	1.66				
0-5 vs. > 10	2.31				
Baseline LDH (ln)	3.09	0.0030	1.64	0.022	4
Localization of disease	1.86	0.0047	1.56	**0.041**	4
Weight	0.99	0.0167	1.75	0.008	1
Gleason Score	1.11	0.2064	1.67	0.016	0
Gleason Score		0.3037	1.64	0.021	0
≤6 vs. 7	1.34				
≤6 vs. ≥8	1.55				
Age	1.03	0.0243	1.86	**0.003**	0
Baseline alkaline phosphatase (ln)	1.49	0.0225	1.81	0.005	0
Baseline hemoglobin (ln)	0.07	0.0079	1.75	0.008	0
Baseline serum PAP (ln)	1.30	0.0038	1.83	0.007	13
PAP positive reaction Immunohistochemistry (< 75% vs. ≥75%)	0.87	0.5464	1.70	0.012	0
ECOG (0 vs. 1)	1.30	0.2766	1.74	0.008	0
Race (Caucasian vs. other)	0.75	0.4601	1.68	0.015	1
Time from diagnosis (yr)	0.99	0.7275	1.70	0.011	0

Thus, three challenges are faced when using Cox PHR model for this study:

a. The treatment effect was not always statistically significant since it depended on the set of covariates chosen and the number of patients excluded from the model;
b. Exclusion of patients from the analysis due to missing covariate data could lead to

biased estimates of the treatment effect;

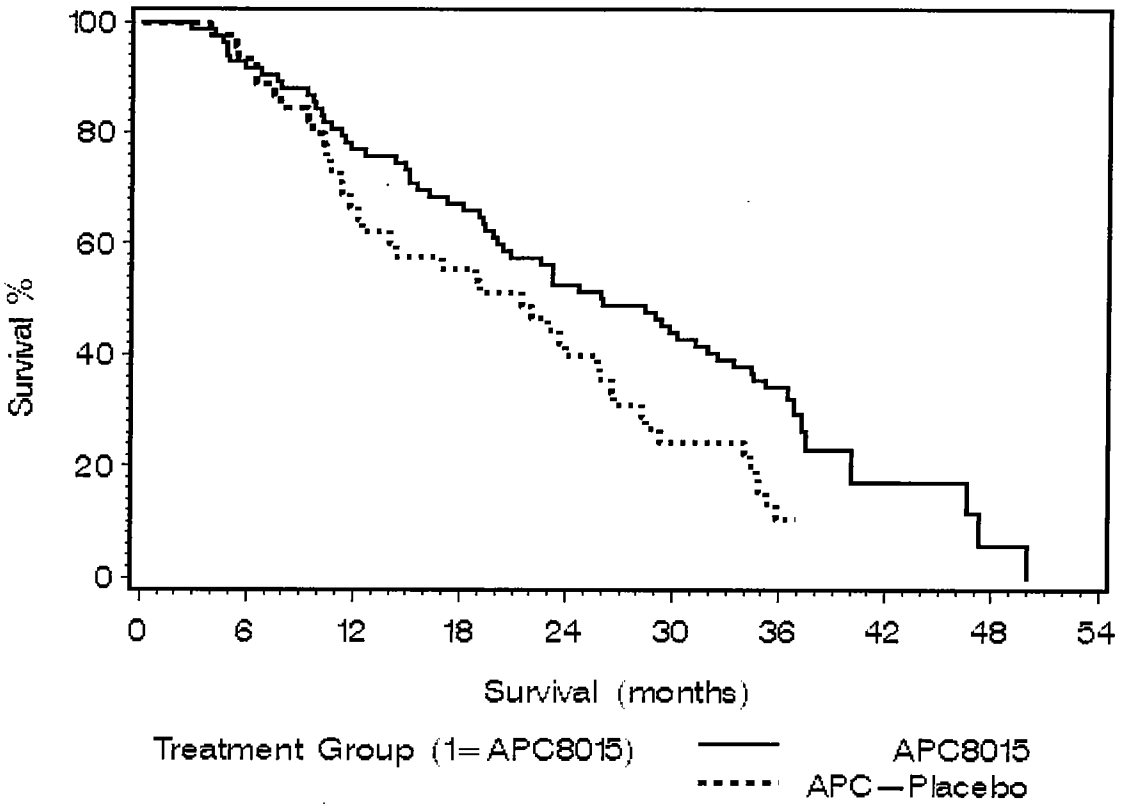
c. Proportion hazards assumption for Cox PHR model may be violated.

Although the difference between the two arms in overall survival was not statistically significant at the level of 0.05 in a few sensitivity analyses, many other sensitivity analyses using Cox PHR model reached statistically significant conclusion at this level. In summary, it appears that the sensitivity analyses support the "statistically significant" finding in overall survival. The quotation marks for statistically significant mean that p-value is less than 0.05 without adjustment for multiple comparisons.

4.1.3 Using all available survival data

The protocol was not designed to collect survival data beyond 36 months. Survival data beyond 36 months were not systematically collected; any additional survival data that Dendreon received from the clinical study centers were not disregarded but included in the D9901 listing entitled "Deaths Known to have Occurred After the 36 Month Follow-up Period" (**refer to Item 8, D9901 CSR, Appendix 16.2.7.7**). Eight patients died after 36 months were included in the analysis as event (all in APC8015 treated group) and the survival curves are presented in the following figure (p =0.011 from log-rank test for comparing the two curves). However, one should interpret the curves after 36-month with caution because of lack of follow-up data for those who are still alive after 36-month.

Figure 4 Kaplan-Meier Plot of Survival Time



4.1.4 Excluding more study centers

Four centers conducted the study under supervision of the same investigator (John Nemunaitis). This reviewer excluded the four study centers (Centers 21, 26, 64, and 68, n=29) and repeated the log-rank test for survival endpoint. The results shows that there are no substantial changes in terms of median survival (23.2 months for APC8015 group and 16.9 months for placebo) and p-values (p=0.01) by excluding 29 subjects from these centers.

4.2 The Primary Efficacy Endpoint

4.2.1 Change of terminology

In the study protocol and its amendments, the primary endpoint was described as **overall time to objective disease progression**. However, the primary efficacy variable in the study report under the BLA submission was presented as **overall time to disease progression**. The reviewers have checked the primary endpoint definitions in protocols against the one in the study report and found that they were the same except using different terminology so it appears that the sponsor did not really change the primary endpoint after completing the study.

4.2.2 P-value change for the primary efficacy endpoint

As reported by the sponsor in Appendix 16.1.9.10 in Study CSR-D9901 of the BLA submission, Study D9901 was unblinded in JUN 2002. At that time, the p-value for time to disease progression in the ITT population was P = 0.088 (log-rank test) when comparing the two arms. The p-value changed to P = 0.085 due to a correction by the independent radiology facility reviewing the scans. A complete clinical audit was performed comparing source documentation at the clinical study centers to the clinical database. Based upon this intensive review of the unblinded data, additional corrections and/or adjustments to the data were warranted, and the p-value changed to P = 0.061 for time to disease progression. A radiographic error was subsequently discovered by the third party CRO responsible for confirmation of objective disease progression, resulting in a p-value change to P = 0.052 for time to disease progression.

The reviewer was unable to duplicate the p-values using the data provided in the submission. After talking to the sponsor, they pointed out that there were a couple typos in the report:

a) Date of last contact for patient 9122-082 should be "30MAR2001" instead of "20MAR2001" as shown in the study report.
b) "1.1 P-value change from p=0.088 to 0.085" as shown in Appendix 16.1.9.10 should be re-written as "1.1 P-value change from p=0.085 to 0.088".

After correcting the typos, the hypothesis tests resulted in:

p =0.085 using blinded review data in June 2002
p =0.088 after the first unblinded review
p =0.061 after the second unblinded review in 2004

p =0.052 after the last unblinded review

It should be noted that the last three p-values (0.088, 0.061, and 0.052) for comparing the curves in time to disease progression (the primary endpoint) were based on unblinded review of selected data. The corrections for errors may or may not be plausible. However, the process of selected corrections in an unblinded fashion could lead to biased estimates of treatment effect. Therefore, the p-value using blinded review data (p=0.085) should be considered as the primary result for the time to disease progression primary endpoint. Others are just exploratory.

When submitting the FDA Statistical Briefing document to the Advisory Committee, the sponsor pointed out that the date for unblinding the data in June 2002 was a mistake in the original submission. The true date should be in July 2002.

4.3 Randomization errors

There were 15 randomization errors. The majority of errors consisted of subjects not being assigned to the randomization slots expected based on the sequence of enrollment. Since Dendreon was blinded to the randomization, the errors were made by the independent third party that did not have access to subject profiles at the time of randomization, and the errors occurred in similar proportions in both treatment arms. An analysis of the potential impact of these randomization errors on overall survival was addressed in a sensitivity analysis by the sponsor in which the treatment assignment of subjects with randomization errors was reversed.

To further evaluate the errors, this reviewer excluded the 15 subjects with randomization errors and repeated the Kaplan-Meier and log-rank analyses, the p-value from testing the difference between the two groups in overall survival was 0.0155. The impact of randomization errors on survival outcome seems to be minimal.

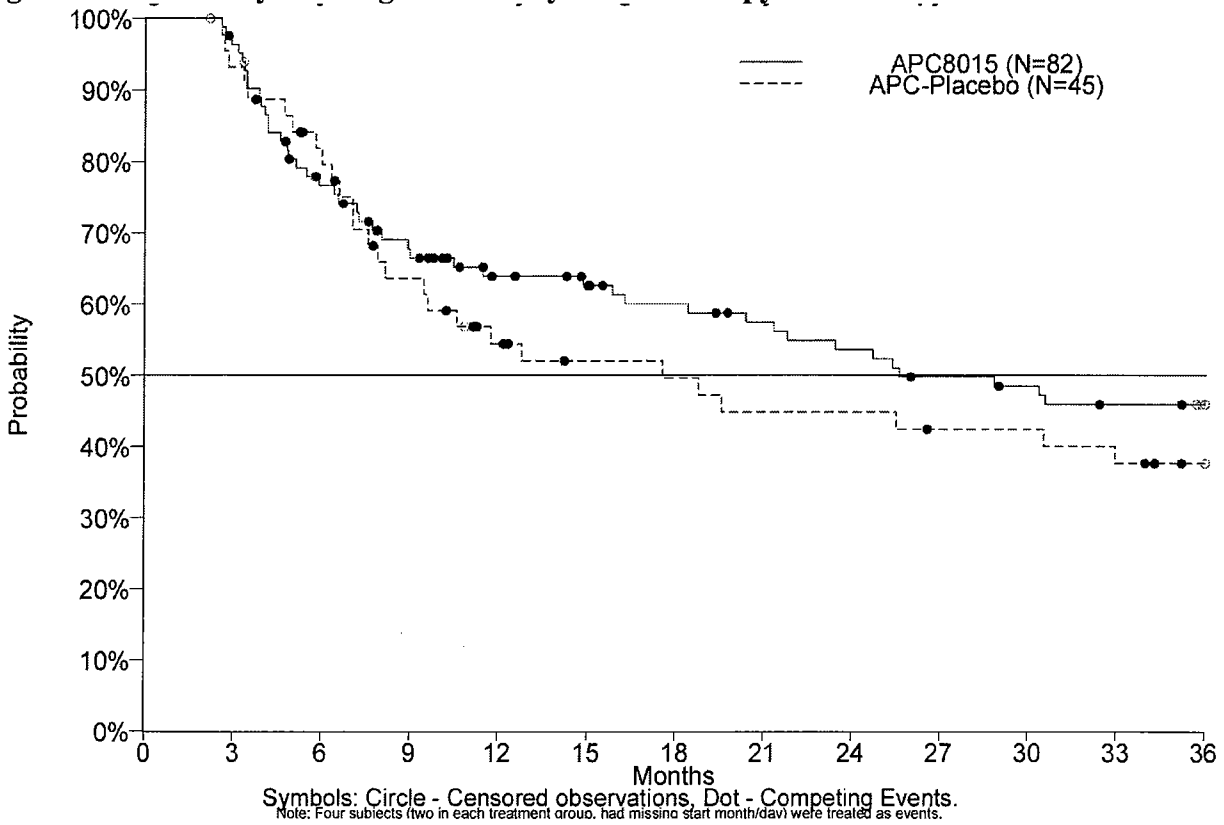
4.4 Use of Chemotherapy during Long-Term Follow-Up

As shown in Table 8, **54.4%** of subjects who received APC8015 were treated with any chemotherapy following disease progression which is numerically lower than **62.8%** of subjects who received APC-Placebo. It is unlikely that the observed better survival among APC8015 treated patients is due to the difference in receiving chemotherapy following disease progression.

However, there is still a concern that the difference between the two arms is due to receiving chemotherapy earlier in patients treated with APC8015 so they have a better survival. To address this concern, further analysis was conducted. The following graph shows probability of being free of any chemotherapy use from randomization. Among those who received chemotherapy approximately before 7 months, there was no obvious difference between the two arms in terms of the timing of receiving chemotherapy. However, the median time to first chemotherapy after randomization was 25.6 months for the APC8015 treat patients, compared to 17.6 months for the placebos. On average, APC8015 treated patients did not received chemotherapy earlier than placebo treated patients. Overall survival analyses also showed that the hazards ratio of placebo

over ACP8015 was greater than one among subjects who received chemotherapy ≤ 7 months after randomization as well as among those who received chemotherapy after 7 months. Similar pattern was observed for the use of docetaxel. It seems that the difference between the two arms is unlikely due to receiving chemotherapy earlier in patients treated with APC8015.

Figure 5 Probability of Being Free of Any Chemotherapy Use from Randomization



4.5 Overall comments

The major efficacy results presented in the efficacy evaluation of the study report for the primary endpoint and overall survival were not analyzed and presented based on a well pre-specified plan. Additional sensitivity analyses conducted by the sponsor and presented in this section showed survival efficacy results with p-values ranging from greater than 0.05 to less than 0.05. Due to the nature of the 'post-hoc analysis', anyone would tend to select any endpoint and/or result for his/her own favor as the primary intentional or unintentionally. Therefore, it is difficult to estimate the type I error rate for this trial and difficult to interpret hypothesis test results for overall survival due to the nature of the analyses.

II. Study D9902

This trial design, patient eligibility, objectives and statistical considerations were the same as those in **Study D9901**. This was a prospective Phase 3, multicenter, double blind, placebo-controlled, randomized trial of immunotherapy with APC8015 for the treatment of subjects with asymptomatic metastatic AIPC. Approximately 120 subjects were planned for enrollment at multiple investigative centers (27 clinical study centers) across the United States. Following a pre-registration and screening process, eligible subjects were randomized to either active treatment or control in a 2:1 ratio (active treatment: control). Following randomization, subjects from both groups underwent a series of 3 standard leukapheresis procedures (in Weeks 0, 2, and 4), and each procedure was followed 2 days later by infusion of either autologous antigen loaded APCs (APC8015; active treatment) or autologous quiescent APCs without antigen (APC-Placebo; control). The treatment phase of the protocol was complete following the third (Week 4) infusion.

This study was initiated shortly after the first study (D9901). After the first study results became available showing no overall significance of TTP, the sponsor did a subset analysis of the first study and found that there was a difference of TTP favoring Provenge arm for subjects who had Gleason score ≤ 7 ($p=0.0014$). At this point, this study had already **enrolled 98 subjects**. The sponsor decided to split the second study into two parts (A and B). Part A (D9902A) contained 98 subjects with all Gleason scores, and part B to enroll subjects only with Gleason score ≤ 7. This BLA contains data from the first part of Study D9902, i.e.: Study D9902A.

1.0 Statistical Analysis Plan

See analysis plan for Study D9901 since the trial design, patient eligibility, objectives and statistical considerations were the same as those in Study D9901.

The original primary and secondary objectives of the study were to compare time to disease progression and time to onset of disease-related pain for subjects treated with APC8015 versus APC-Placebo. Following discussions with the FDA and prior to performing any analyses, the statistical analysis plan was revised to elevate overall survival to a secondary endpoint.

Subjects were followed for survival for 3 years following their date of randomization. Subjects who were alive at 3 years following randomization were censored at 3 years from their date of randomization. Per the statistical analysis plan, survival was evaluated after the 96th subject had been followed for 3 years following randomization, at which point 2 additional subjects were censored for survival and included in the 36-month statistic. **The sponsor stated that Cox PHR model with covariates were using the same from Study D9901 before they looked at the data.**

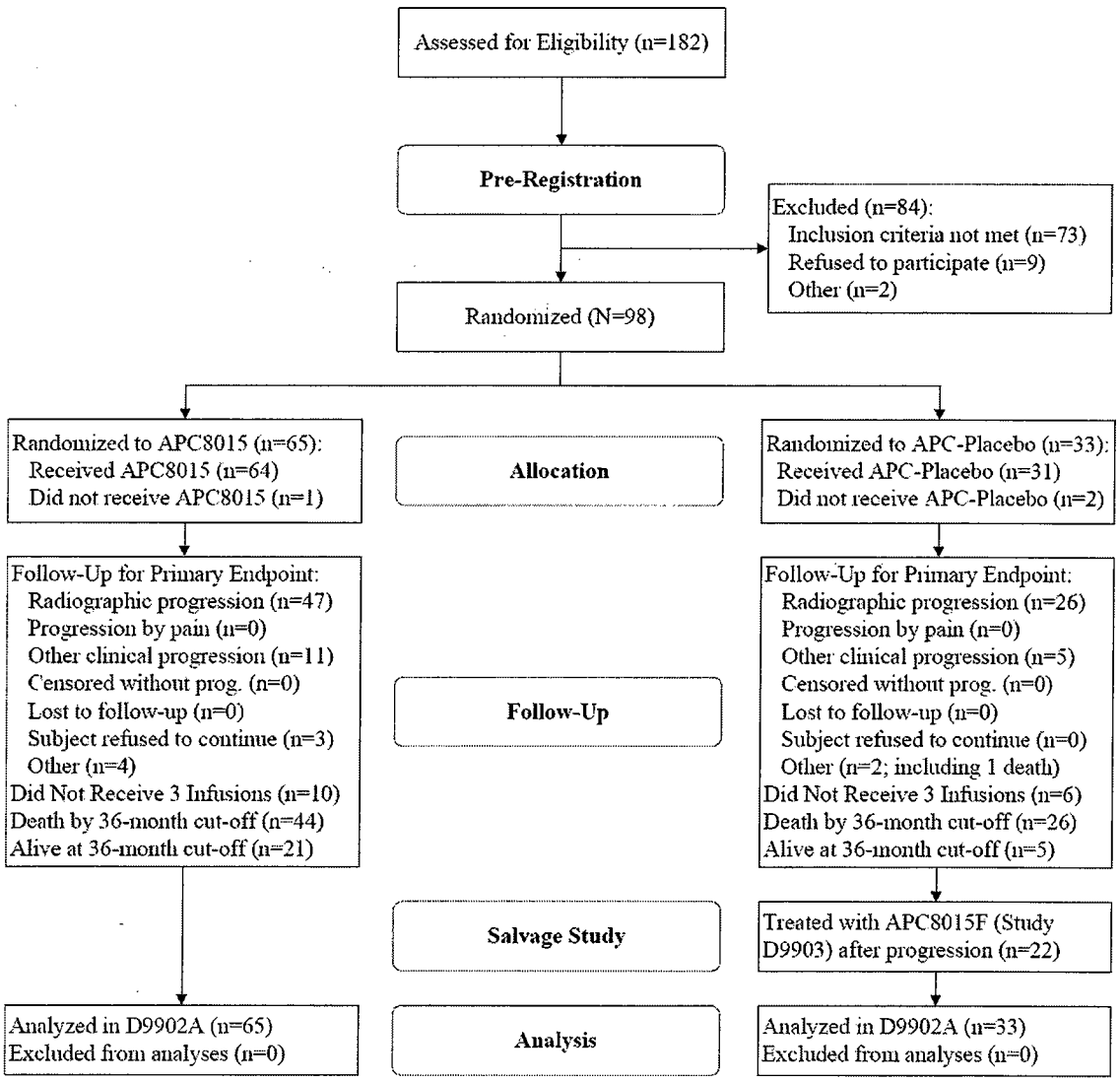
These data were analyzed using the ITT and Safety populations. The ITT population included all randomized subjects (n = 98), and the Safety population included all subjects who underwent at least 1 leukapheresis (n = 96).

2.0 Efficacy Evaluation

2.1 Disposition of subjects

Of the 182 subjects screened for eligibility, 98 subjects were randomized between **12 MAY 2000 and 21 MAR 2003**. Of these, 65 subjects were randomized to receive APC8015 and 33 subjects were randomized to receive APC-Placebo. A total of 96 subjects underwent at least one leukapheresis procedure and a total of 95 subjects received at least one infusion. Figure 6 presents a schematic of the subject disposition summary for this trial.

Figure 6 Schematic of Subject Disposition



For all randomized subjects, 28 of 98 (28.6%) subjects had protocol deviations with similar

percentages of subjects in both treatment groups reporting these deviations (17 of 65 [26.2%] subjects treated with APC8015 compared with 11 of 33 [33.3%] subjects treated with APC-Placebo). Major protocol eligibility deviations were issued for 2 subjects (1 subject treated with APC8015 and 1 subject treated with APC-Placebo). The subject treated with APC8015 had no evidence of metastatic disease at entry and entered the study with a baseline testosterone value of > 50 ng/mL. The subject treated with APC-Placebo was not medically or surgically castrate at study entry. The remainders of the deviations were either minor eligibility deviations or non-eligibility related.

Based on a review of the randomized treatment assignments performed by Dendreon after the study was unblinded, it was determined that **18 randomization errors occurred**. The majority of errors consisted of subjects not being assigned to the randomization slots expected based on the sequence of enrollment. Since Dendreon was blinded to the randomization, the errors were made by the independent third party that did not have access to subject profiles at the time of randomization, and the errors occurred in similar proportions in both treatment arms.

2.2 Demographics and other baseline characteristics

A summary of demographics and baseline characteristics is provided in Table 12. The demographic characteristics were similar between the 2 treatment groups. All subjects enrolled in this trial were male, and the majority of subjects were Caucasian (91.8%). The median age in this population was 71.0 years; ages ranged from 51 years to 87 years. The majority of subjects from both treatment groups had a baseline ECOG performance status of 0 (78.5% of subjects treated with APC8015 and 69.7% of subjects treated with APC-Placebo).

Table 12 Summary of Subject Demographics and Baseline Characteristics, ITT

Parameter	APC8015 (n = 65)	APC-Placebo (n = 33)	Total (N = 98[a])	p-value[b]
Age (years)				0.544
N	65	33	98	
Mean	69.6	70.6	69.9	
SD	8.2	7.8	8.1	
Median	70.0	71.0	71.0	
Minimum	51.0	57.0	51.0	
Maximum	84.0	87.0	87.0	
Race, n (%)				0.474
Caucasian	59 (90.8)	31 (93.9)	90 (91.8)	
African American	2 (3.1)	2 (6.1)	4 (4.1)	
Hispanic	1 (1.5)	0 (0.0)	1 (1.0)	
Unknown	3 (4.6)	0 (0.0)	3 (3.1)	
Weight (lbs)				0.958
N	62	32	94	
Mean	195.7	195.3	195.6	
SD	33.0	36.8	34.2	
Median	191.3	184.0	190.0	
Minimum	131	145	131	
Maximum	283.5	276.5	283.5	
ECOG Performance Status, n (%)				0.343
0	51 (78.5)	23 (69.7)	74 (75.5)	
1	14 (21.5)	10 (30.3)	24 (24.5)	
Serum PSA (ng/mL)				0.668
N	64	33	97	
Mean	153.7	177.1	161.6	
SD	217.2	313.8	252.8	
Median	61.3	44.0	53.3	

Parameter	APC8015 (n = 65)	APC-Placebo (n = 33)	Total (N = 98[a])	p-value[b]
Minimum	8.0	8.2	8.0	
Maximum	936.5	1342.5	1342.5	
Serum PAP (ng/mL)				0.277
N	55	31	86	
Mean	23.8	13.7	20.1	
SD	47.1	26.7	41.1	
Median	4.5	5.1	5.1	
Minimum	0.7	0.6	0.6	
Maximum	230.0	144.0	230.0	
Alkaline Phosphatase (U/L)				0.263
N	65	33	98	
Mean	285.7	167.2	245.8	
SD	591.3	166.6	493.0	
Median	140.0	105.0	116.0	
Minimum	50.0	34.0	34.0	
Maximum	3900.0	923.0	3900.0	
LDH (U/L)				0.893
N	63	33	96	
Mean	248.6	242.8	246.6	
SD	219.2	161.3	200.3	
Median	187.0	179.0	185.5	
Minimum	101	116	101	
Maximum	1730.0	730.0	1730.0	
Hemoglobin (g/dL)				0.889
N	65	33	98	
Mean	12.7	12.8	12.8	
SD	1.5	1.5	1.5	

Parameter	APC8015 (n = 65)	APC-Placebo (n = 33)	Total (N = 98[a])	p-value[b]
Median	12.8	12.6	12.8	
Minimum	9.2	9.0	9.0	
Maximum	15.8	15.3	15.8	
Time from Diagnosis to Randomization (weeks)				0.460
N	65	33	98	
Median	288.1	326.0	291.4	
Minimum	71	54	54	
Maximum	669	590	669	

[a] Unknown values were not categorized; therefore, all parameters do not represent 98 subjects.

[b] p-value compares treatment groups using a Cochran-Mantel-Haenszel test controlled by center for categorical data and a 2-way analysis of variance with factors of treatment group and center for continuous data.

The PAP immunohistochemistry results are summarized in Table 13. A higher percentage of subjects in the APC8015 group than in the APC-Placebo group had tumor specimens with at least 75% PAP-positive cells

Table 13 Summary of PAP Immunohistochemistry, ITT

	APC8015 (n = 63)	APC-Placebo (n = 33)	Total (N = 96)	p-value[a]
Percent PAP Positive, n (%)				0.004
< 25%	0 (0.0)	1 (3.0)	1 (1.0)	
25% - 74%	17 (27.0)	19 (57.6)	36 (37.5)	
≥ 75%	46 (73.0)	13 (39.4)	59 (61.5)	

[a] p-value compares treatment groups using a Cochran-Mantel-Haenszel test.

For this study, 85 of 98 subjects (86.7%) had a Gleason score assigned by a central pathology laboratory prior to randomization. The Gleason scores obtained by local pathology facilities were used for those subjects not given Gleason scores by the central pathology facility. A Gleason score was unattainable for 1 subject (randomized to APC8015). A summary of subjects' Gleason score and tumor status at baseline is presented in Table 14. Overall, a majority of the subjects had a Gleason score ≤ 7 (61 subjects [62.9%] versus a Gleason score ≥ 8 (36 subjects [37.1%]).

Table 14 Summary of Gleason Score and Tumor Status at Baseline, ITT

	APC8015	APC-Placebo	Total	p-value[a]
Gleason Score				
N	64	33	97	
Median	7	7	7	
Minimum	5	4	4	
Maximum	9	10	10	
Score. n (%)				0.423
≤ 6	15 (23.4)	9 (27.3)	24 (24.7)	
= 7	29 (45.3)	8 (24.2)	37 (38.1)	
≥ 8	20 (31.3)	16 (48.5)	36 (37.1)	
Localization of Disease				0.221
N	65	33	98	
Bone metastases only	31 (47.7)	10 (30.3)	41 (41.8)	
Soft tissue metastasis/pelvis recurrence only	7 (10.8)	7 (21.2)	14 (14.3)	
Both bone metastases and soft tissue metastasis/pelvic recurrence	27 (41.5)	16 (48.5)	43 (43.9)	
Number of Bone Metastases per subject[b]	61	32	93	0.078
0	5 (8.2)	7 (21.9)	12 (12.9)	
1 – 5	19 (31.1)	11 (34.4)	30 (32.3)	
6 – 10	6 (9.8)	2 (6.3)	8 (8.6)	
> 10	31 (50.8)	12 (37.5)	43 (46.2)	

[a] p-value compares treatment groups using a Cochran-Mantel-Haenszel test controlled by center for categorical data and a 2-way analysis of variance with factors of treatment group and center for continuous data.

[b] Five subjects did not have readings of baseline bone scans by the central radiology facility.

It should be noted that p-values provided by the sponsor in the above tables should not be considered as those from a hypothesis test for the difference between the two arms. They just reflect the chance of obtaining the observed difference (and more extreme) between the two arms in these demographic and baseline characteristics factors when in fact the two samples were randomly drawn from the same population.

2.3 Efficacy Results

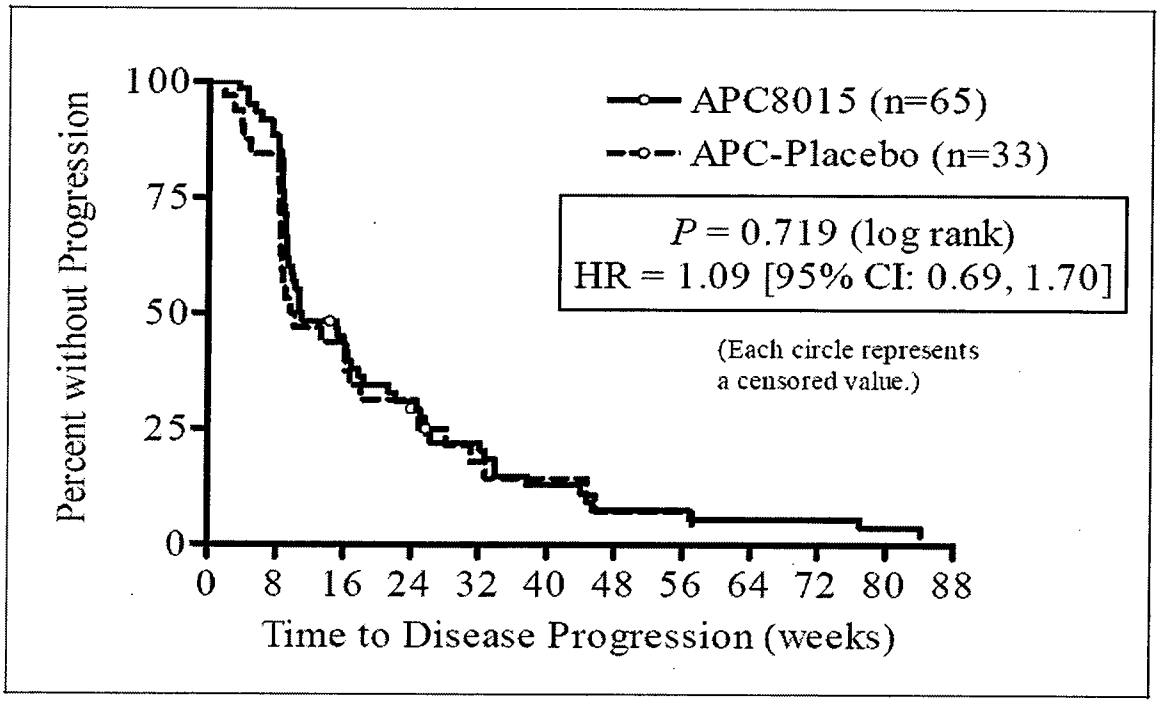
2.3.1 The primary endpoint

Out of 98 subjects randomized on this study, 89 subjects (90.8%) contributed a progression event. Seventy-three subjects (74.5%) had progression documented by serial imaging (as determined by independent blinded radiology review) and 16 subjects (16.3%) had progression based on clinical events other than radiographic events, as defined in the protocol. Of the 73

subjects who progressed based on imaging studies, 36 subjects treated with APC8015 and 17 subjects treated with APC-Placebo progressed based on bone disease, while 11 subjects treated with APC8015 and 9 subjects treated with APC-Placebo progressed based on soft-tissue disease.

When the Kaplan-Meier curves for overall time to disease progression were compared, there was no significant difference between the 2 arms (P = 0.719, log rank; unadjusted HR = 1.09 [95% CI: 0.69, 1.70]; Figure 7). The estimated median time to disease progression was 10.9 weeks in the APC8015 group compared with 9.9 weeks in the APC-Placebo group.

Figure 7 Primary Efficacy Endpoint, Overall Time to Disease Progression (Kaplan-Meier Method), ITT



2.3.2 Overall survival

The analysis of the 3-year survival data was based on the ITT population of all 98 randomized subjects. As described in the protocol and statistical analysis plan, every subject was followed until death or the pre-specified cut-off of 36 months. Per the statistical analysis plan, survival was evaluated after the 96th subject had been followed for 3 years following randomization, at which point 2 additional subjects were censored for survival and included in the 36-month analysis.

As shown in Figure 8, the unadjusted HR was 1.27 ([95% CI: 0.78, 2.07]; P = 0.331, log rank). The median survival time for subjects treated with APC8015 was 3.3 months longer than that for subjects treated with APC-Placebo (median survival times of 19.0 months [95% CI: 13.6, 31.9] and 15.7 months [95% CI: 12.8, 25.4], respectively). Table 15 and Table 16 present key summary statistics that characterize the differences between the two treatment arms. Overall, 28 subjects were alive at the 36-month follow-up visit and had survival times that were censored at that time

(21 subjects treated with APC8015 and 7 subjects treated with APC-Placebo). At the end of the study, the proportion of subjects in the APC8015 group who were alive was 50% higher than that of the APC-Placebo group (31.6% versus 21.2%, respectively).

Figure 8 Overall Survival (Kaplan-Meier Method), ITT

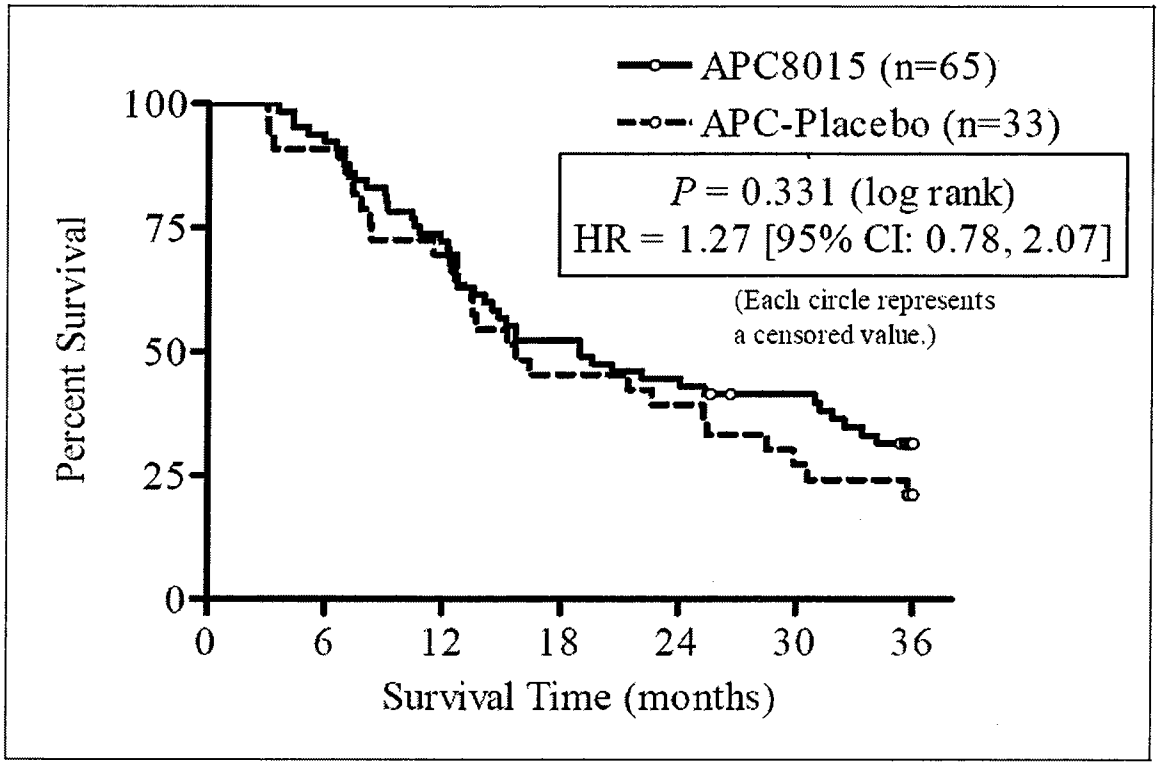


Table 15 Overall Survival Summary Statistics, ITT

Treatment	Number of Subjects	Deaths Over 36-Month Follow-Up	Subjects Alive at 36-Month Visit	Death Percentiles, months		
				25%	50%	75%
APC8015	65	44 (67.7%)	21 (32.3%)	10.8	19.0	≥31.9
APC-Placebo	33	26 (78.8%)	7 (21.2%)	8.3	15.7	30.6

Table 16 Kaplan-Meier Survival Rate Estimates, Percent ITT

Treatment	6 months	12 months	18 months	24 months	30 months	36 months
APC8015	92.3	72.3	52.3	44.6	41.5	31.6
APC-Placebo	90.9	69.7	45.5	39.4	27.3	21.2

A PHR model was developed to test the robustness of the survival results in the randomized, placebo controlled companion study, D9901. The 5 baseline prognostic factors described in this model were LDH (natural log [ln]), PSA (ln), localization of disease, number of bone metastases, and body weight (lbs). When the model with these prognostic factors identified in D9901 was applied to D9902A, localization of disease, LDH, and weight were no longer significant, but PSA and lesion count continued to be strong predictors of survival. In this multivariate analysis, the treatment effect was significant (adjusted HR = 1.92; P = 0.023, Table 17). In the model below, the HR for treatment with APC8015 indicates that the death rate of subjects randomized to receive APC-Placebo is 1.92-fold higher than that for subjects randomized to receive APC8015.

Table 17 Proportional Hazards Regression Model of Survival

		95% CI for HR		
	HR	**Lower**	**Upper**	**p-value**
Treatment with APC8015	1.92	1.09	3.35	0.023
Baseline PSA (ln)	1.35	1.09	1.67	0.006
Baseline LDH (ln)	0.96	0.53	1.74	0.885
Localization of Disease	1.67	0.95	2.92	0.075
Number of Bone Metastases				0.002
Lesion count: 0-5 lesions versus 6-10 lesions	3.67	1.28	10.50	
Lesion count: 0-5 lesions versus > 10 lesions	2.95	1.58	5.50	
Weight	1.00	0.99	1.00	0.263

N = 88: Events = 62; Censored = 26. and cases with missing values = 10

A HR greater than 1.0 indicates a benefit in favor of APC8015.

In order to assess whether the use of chemotherapy, in particular the use of docetaxel, influenced the survival results, the two treatment groups were compared with respect to the administration of chemotherapy following protocol treatment. This analysis revealed that there was not a statistically significant imbalance in the administration of chemotherapies between treatment groups. Based on data from 93 of 98 randomized subjects, 66.7% of subjects who received APC8015 and 54.5% of subjects who received APC-Placebo went on to be treated with any chemotherapy. More specifically, 56.1% of subjects who received APC8015 and 40.6% of subjects who received APC-Placebo were subsequently treated with taxane-based chemotherapy; 38.6% of APC8015-treated subjects and 34.4% of APC-Placebo-treated subjects received docetaxel. No statistically significant differences at the 0.05 level in chemotherapy use were found. (Only the type of chemotherapy and the date of initiation were collected for this trial. No data on the dose or duration of chemotherapy are available.)

3.0 Safety Evaluation

Ninety-six of 98 subjects randomized on this trial underwent at least 1 leukapheresis procedure and are included in the safety population. As indicated in the protocol, 3 infusions were to be administered over a period of 4 weeks. A total of 82 of 95 treated subjects (86.3%) received all 3 infusions. Ten of 95 treated subjects (10.5%) received 2 infusions, 7 subjects treated with APC8015 and 3 subjects treated with APC-Placebo. Three of 95 treated subjects (3.2%) received only 1 infusion during the study, 2 subjects treated with APC8015 and 1 subject treated with APC-Placebo.

A total of 82 of 95 subjects (86.3%) in both treatment groups who received at least 1 infusion received all 3 infusions, and no subject was withdrawn from the trial due to unacceptable toxicity.

The majority of subjects in this study experienced mild to moderate AEs. Events that occurred most frequently were generally mild and resolved within 1 to 3 days.

Grade 3 or 4 AEs were observed in a total of 20 subjects (30.8%) in the APC8015 treatment and 8 subjects (25.8%) in the APC-Placebo treatment arm. Only 7 subjects experienced SAEs that were considered related to study treatment, 6 subjects treated with APC8015 and 1 subject treated with APC-Placebo.

A total of 70 deaths occurred within the 3-year reporting period, with a lower percentage of deaths in the APC8015 group versus the APC-Placebo group (67.7% deaths [44 of 65 subjects] versus 78.8% [26 of 33 subjects], respectively). One subject treated with APC8015 a Grade 5 event of cerebrovascular accident and 1 subject treated with APC-Placebo a Grade 5 event of small cell lung cancer; both of these events occurred during the long-term portion of the trial (5 and 6 months, respectively, after disease progression). One additional subject treated with APC-Placebo died of cardiac failure prior to disease progression but approximately 6 months after his last infusion of APC-Placebo. No deaths occurred within 30 days of product infusion.

Detailed review on safety can be seen in the medical reviewer's memo.

4.0 Statistical Findings and Comments

Similar issues identified in Study 9901 will be not discussed again in this section. These issues include the inflation of type I error rate, change of terminology for the primary endpoint, and randomization errors. However, results from the sensitivity analyses for overall survival in Study D9902A are quite different from those for Study D9901. There are some other unique issues for D9902A and they are discussed together with overall survival in the following.

4.1 Additional sensitivity analyses for overall survival using Cox model

The log-rank test for treatment effect on overall survival resulted in a p-value of 0.331 and the p-value was reduced to 0.023 when adjusting for a set of covariates using Cox proportional hazards model as reported by the sponsor [Cox model (I) in Table 18]. The results were verified by the statistical reviewer.

When conducting additional sensitivity analyses with different set of covariates in Cox model, most of them did not reach statistical significance for the difference between the two groups. The reviewer also evaluated all demographics and baseline characteristics parameters and their impacts on overall survival. The treatment effect was NOT statistically significant when adjusting for any individual covariate in Cox PHR model.

Table 18 Treatment effect on overall survival using Cox model with different set of covariates (total subjects: 98)

Testing Method	Hazard Ratio	P-value	Covariate Adjustment	#Patients Missing
Log-rank test*	1.27	0.331	None	0
Cox Model (I)*	1.92	0.023	Localization of disease, PSA(ln), LDH(ln), Weight, #Bone metastases (≤5, 6-10, >10) [Strata: bisphosphonate]	10
Cox Model (II)~	1.49	0.122	Localization of disease, Gleason Score (≤6, 7, ≥8), PSA (<20, 20 - <100, ≥100)	2
Cox Model (III)~	1.12	0.642	Age, LDH(ln)	2

*: The sponsor's analyses and confirmed by the reviewer
~: Reviewer's analysis

Similar to Study 9901, 10 subjects were excluded from the analysis due to missing data in some of the covariates when Cox model (I) was used by the sponsor (resulting in a reduction of p-value from 0.33 to 0.023). Therefore, the statistical reviewer further analyzed the survival data by breaking subjects into two groups: 1) subjects with any missing covariate data and were excluded from the analysis in Cox model (I) by the sponsor; 2) subjects with complete covariates data for Cox model (I). As presented in Table 19, one can see that patients without complete covariate data had poor median survival in APC8015 treated group, compared to the rest of the APC0805

treated patients. In contrast, placebo patients with missing covariate data had longer median survival relative to the rest of the placebo patients. The randomization principle may be seriously compromised when using Cox model (I) and exclusion of subjects can lead to biased estimates of treatment effects. Thus, the results from any analyses with exclusion of subjects should be interpreted cautiously. Similar to Study D9901, if imputing missing covariate data in order to include all randomized patients in the analysis using Cox Model (I), the difference between the two arms in overall survival could be statistically significant or non-statistically significant depending on what imputation approach was used.

Table 19 Impact of missing covariate data in Cox Model (I)

Complete Covariate Data	APC8015			Placebo		
	N	# Deaths	Median Survival (mons)	N	# Deaths	Median Survival (mons)
No	8	6	13.2	2	2	20.9
Yes	57	38	19.6	31	24	15.3

In summary, there is not enough evidence in support of the sponsor's finding of statistically significant difference between the two arms in overall survival using Cox model (I) though there was a trend of difference in overall survival in Study D9902A.

4.2 Discrepancy between D9901 and D9902A in Cox model (I)

The sponsor stated that they planned to use the same set of covariates in Cox model (I) from Study D9901 for Study D9902A before they looked at the data in order to support their statement that the analysis using Cox model was pre-specified in Study D9902A. However, there is a discrepancy between the two studies in terms of using Cox PHR model.

As shown in the randomization plan in both Phase III studies, subjects were stratified by clinical study center and use of bisphosphonate therapy (yes or no) prior to being randomized. The Cox PHR model the sponsor used for Study D9901 did not incorporate use of bisphosphonate therapy as strata in the model while the Cox model for Study D9902A did. Therefore, the two models are different. When analyzing data from Study D9902A using the exact same Cox model in Study D9901 (without use of bisphosphonate therapy as strata in the model), hazards ratio changed from 1.92 to 1.79 and p-value for comparing the two treatment groups changed from 0.023 to 0.036.

Given the fact that data were available at the time when the sponsor developed an analysis plan for survival for Study D9902A, there was a chance that the analysis plan might be developed under the influence with the knowledge of the survival data intentionally or unintentionally.

4.3 Inappropriate presentation of data

The sponsor presented overall survival data for Study D9902A as below (Table 15 in this review memo). Reporting 21 subjects (APC8015 treated) alive in Column "**Subjects alive at 36-month visit**" is misleading. Two subjects who censored before 36-month (alive at the end of follow-up before 36-month) were treated as alive at 36-month. In fact, there were only 19 subjects in APC8015 group who were known to be alive at 36-month. Thus, number of subjects alive a 36-month in APC8015 group should be estimated in the range of 19-21 (29.2%-32.3%).

Table 10 Overall Survival Summary Statistics, ITT

Treatment	Number of Subjects	Deaths Over 36-Month Follow-Up	Subjects Alive at 36-Month Visit	Death Percentiles, months		
				25%	50%	75%
APC8015	65	44 (67.7%)	21 (32.3%)	10.8	19.0	≥31.9
APC-Placebo	33	26 (78.8%)	7 (21.2%)	8.3	15.7	30.6

Source: Post-Text Table 14.1.1 (16 MAY 2006); Post-Text Table 14.2.6 (01 JUN 2006)

4.4 Failure of duplicating the results for patients with Gleason score ≤ 7

Study 9902 was initiated shortly after the first study (D9901). After the first study results became available showing no overall significance of TTP, the sponsor did a subset analysis of the first study and found that there was a difference of TTP favoring APC8015 treated arm for subjects who had Gleason score ≤ 7 (p=0.0014). The sponsor decided to split the second study into two parts (A and B). Part A (D9902A) contained 98 subjects with all Gleason scores, and part B to enroll subjects only with Gleason score ≤ 7. When the survival data available later for D9901, median overall survival was 28.7 months for APC8015 treated subjects with Gleason ≤ 7 and 21.4 months for placebo subjects (p=0.0218). However, after looking at the data in Study D9902A, the results could not be repeated. Among patients with Gleason score ≤ 7 in Study D9902A, the median TTT was 10.7 weeks for APC8015 treated subjects and 14.7 weeks for placebo subjects (p=0.6024). Median overall survival was 20.1 months for APC8015 treated subjects and 21.4 months for placebo subjects (p=0.7293). Numerically, APC8015 treated patients did worse than the placebo in the subgroup of patients with Gleason score ≤ 7 in Study D9902A. The discrepancy implies that post-hoc analysis results should be, in general, treated as proof-of-concept and need to be confirmed with other study.

III. Integrated Summary and Other Findings

1.0 Summary of Efficacy

The IND application for APC8015 was submitted in November 1996 for the treatment of prostate cancer. Early Phase 1 and 2 clinical studies in men with androgen independent prostate cancer (AIPC) were conducted to test the safety and preliminary efficacy of APC8015. The results demonstrated the following: 1) Intravenous infusions of APC8015 in subjects with prostate cancer were generally well tolerated with no dose limiting toxicities observed; 2) Prostate-specific antigen (PSA) reductions of >50% in approximately 10% of subjects were noted, as well as one striking objective response; 3) Three doses of APC8015 resulted in substantial PA2024-specific immune responses and appeared to delay time to disease progression compared to historic controls.

The Phase 3 clinical development program was initiated in 2000 and was designed to evaluate the safety and effectiveness of APC8015 to address the unmet medical need of treating men with asymptomatic metastatic AIPC. At the time the Phase 3 program was designed, there were no approved therapies for men with asymptomatic or minimally symptomatic AIPC. Standard and experimental chemotherapies available at the time were palliative and associated with significant toxicity. It was not until May 2004 that docetaxel was approved and shown to prolong survival in AIPC. However, there remains a clear need for less toxic therapies that benefit patients with AIPC, particularly those without cancer-related pain, many of whom elect not to pursue chemotherapy.

The pooled survival data are derived from all 225 randomized subjects in D9901 and D9902A. A total of 223 subjects were followed until death or until a cut-off of 36 months after their date of randomization. Two additional subjects in D9902A were censored at **25.6** months and **26.7** months as a result of a pre-defined data cut-off point.

The summary of post-hoc analyses on the overall survival endpoint is presented in Figure 9 and Table 20 (P = 0.011, log rank, stratified by study). The HR for treatment, based on the Cox PHR model with no covariates, was 1.50, indicating a 33% reduction in the death rate for subjects treated with APC8015 compared to APC-Placebo. The median survival time for subjects treated with APC8015 was 4.3 months longer than that for subjects treated with APC-Placebo (median survival times of 23.2 months and 18.9 months, respectively).

Figure 9 Integrated D9901 and D9902A Overall Survival (Kaplan-Meier Method), ITT

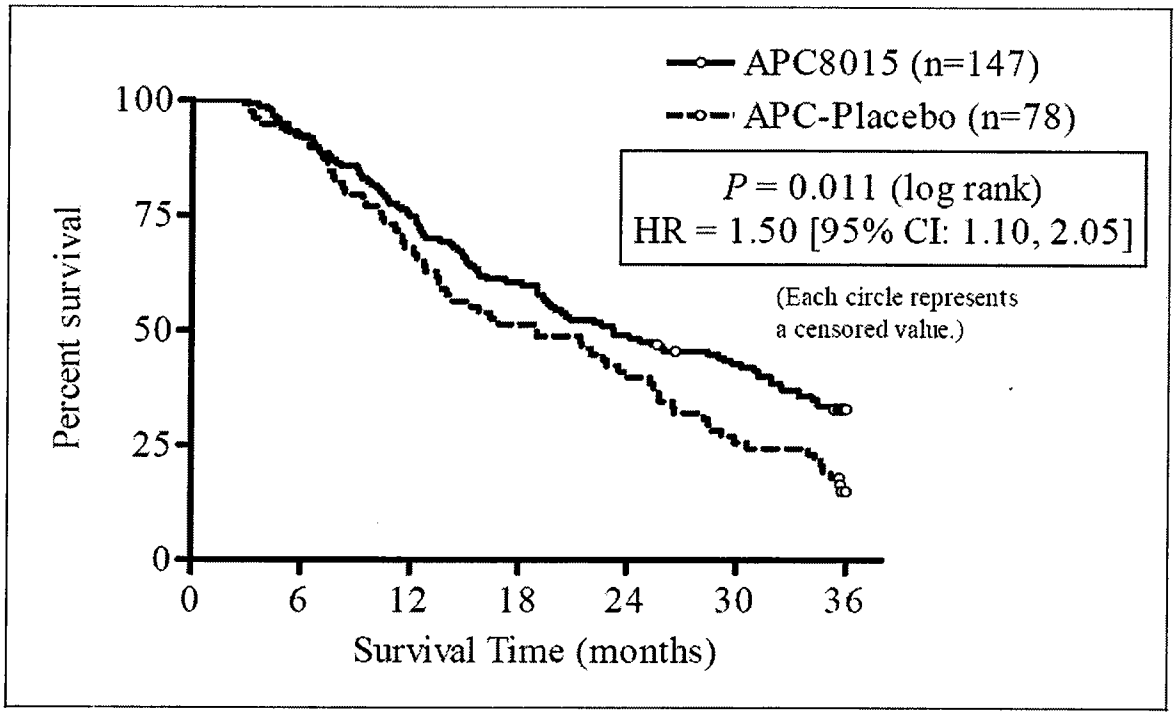


Table 20 Summary of Overall Survival in Studies D9901 and D9902A

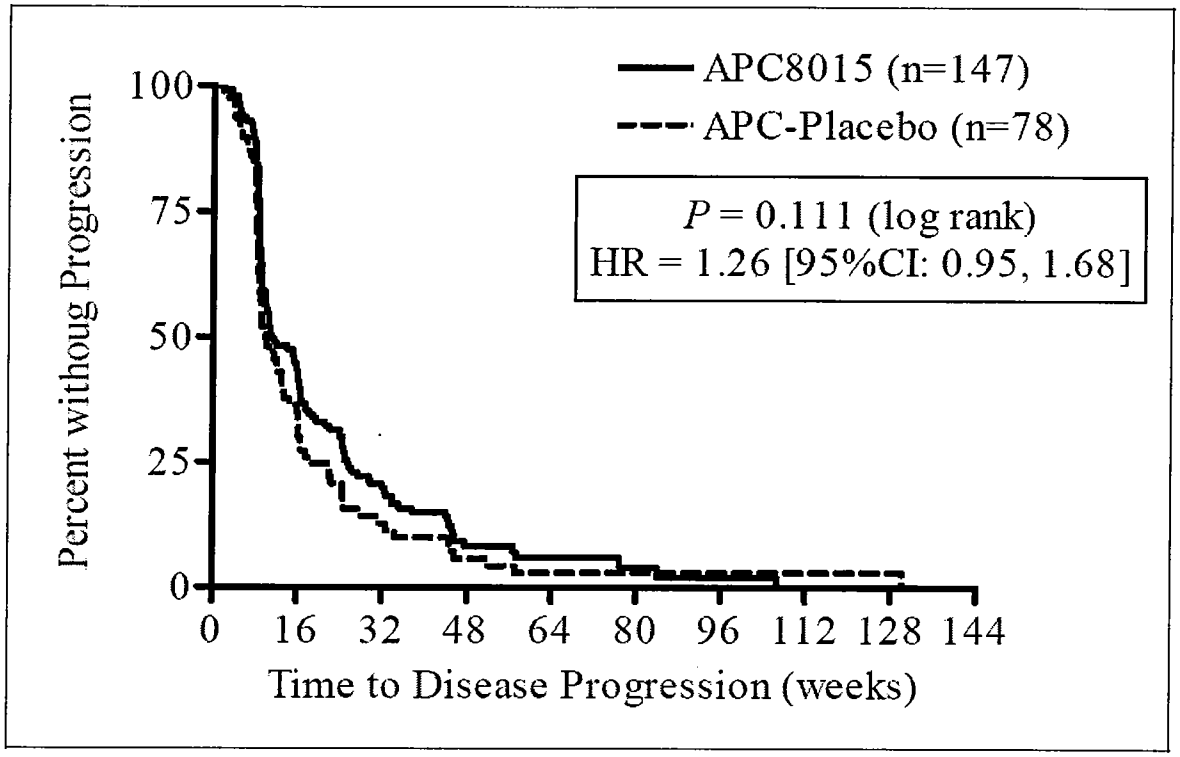
	Study D9901 Months (95% CI)		Study D9902A Months (95% CI)		Integrated Months (95% CI)	
	APC8015 N=82	APC-Placebo N=45	APC8015 N=65	APC-Placebo N=33	APC8015 N=225	APC-Placebo N=78
Median Survival	25.9 (20.0, 32.4)	21.4 (12.3, 25.8)	19.0 (13.6, 31.9)	15.7 (12.8, 25.4)	23.2 (19.0, 31.0)	18.9 (13.5, 25.3)
Hazard Ratio[a]	1.71 (1.13, 2.58)		1.27 (0.78, 2.07)		1.50 (1.10, 2.05)	
Overall Survival (log rank test)	$P = 0.010$		$P = 0.331$		$P = 0.011$	

[a] A hazard ratio greater than 1.0 favors APC8015.

Abbreviations: CI = confidence interval.

When the Kaplan-Meier curves for overall time to disease progression were compared, there was no significant difference between the 2 treatment arms in the integrated D9901 and D9902A ITT population (P = 0.111, log rank; HR = 1.26 [95% CI: 0.95, 1.68]; Figure 10). The estimated median time to disease progression was 11.1 weeks in the APC8015 group compared with 10.0 weeks in the APC-Placebo group.

Figure 10 Integrated D9901 and D9902A Overall Time to Disease Progression (Kaplan- Meier Method), ITT



2.0 Summary of Safety

Subjects who underwent at least 1 leukapheresis procedure were included in the safety population. The safety population consists of **669 subjects** who underwent at least 1 leukapheresis procedure and were to be treated with APC8015, APC-Placebo, APC-Placebo followed by APC8015F, or APC8026 (a product similar to APC8015 but with a different method of manufacture and cell composition).

Detailed review on safety can be seen in the clinical review's memo.

3.0 Statistical Findings and Comments

3.1 Inconsistent efficacy results between the two trials

The study design was identical in the two studies and the outcome for the primary endpoint, time to disease progression, was quite similar between the two studies. As shown in the following table, the median time to disease progression for placebo patients were 9.1 weeks and 9.9 weeks for Studies D9901 and D9902A, respectively; the median time to disease progression for APC8015 treated patients were 11.0 weeks and 10.9 weeks for Studies D9901 and D9902A, respectively.

However, the survival experience was quite different for the two studies. As shown in Table 21, the median survival for placebos was 15.7 months in Study D9902A, but 21.4 months in Study D9901. The median survival for APC8015 treated patients in Study D9902A was 19 month which is shorter than that for the placebo patients in Study D9901. This could be due to the baseline imbalance between the two studies, but could also due to sampling variation.

Table 21 Comparisons of TTP and overall survival between the two Phase III studies

Study	Median Time to Progression (wks)		Median Survival Time (Months)	
	APC8015	Placebo	APC8015	Placebo
D9901*	11.0	9.1	25.9	21.4
D9902A	10.9	9.9	19.0	15.7

* using the blinded review data

3.2 Strength of overall efficacy evidence

The Guidance for Industry--Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products (see ref: [1]) explicitly says "*With regard to quantity, it has been FDA's position that Congress generally intended to require at least two adequate and well-controlled studies, each convincing on its own, to establish effectiveness.*" It implies that positive results from at least two adequate and well-controlled trials are, in general, required for licensure application. Two positive trials not only substantially reduce the chance of making a false claim, but also assure that a result could be more robust, replicated in different settings, and more representative to the targeted population.

In each adequate, well-controlled (pivotal) trial, the probability of erroneously rejecting the null hypothesis (significance level) needs to be pre-specified. Conventionally, this level is set at 0.05 (two-sided). If two trials are required to show statistically significant on the primary endpoint, the operational type I error rate for approving an ineffective product should be controlled under a level of 1/1600 (=1/40 × 1/40), from a statistical perspecitve. This is because taking the

operational view, we simply regard a two-sided test at the 0.05 level as being a one-sided test at the 0.025 level (1/40). (see refs: Senn[2], Fisher[3])

The key efficacy evidence in this BLA is the difference between the two arms in overall survival. The log-rank tests for comparing the two arms in overall survival resulted in a p-value of 0.01 and 0.331 for Studies D9901 and D9902A, respectively. As shown in the integrated analyses, analyzing the pooled survival data derived from all 225 randomized subjects in D9901 and D9902A showed a difference between the two arms in overall survival with a p-value of 0.011. Although analyzing the pooled data from the two studies supported the "statistically significant" finding in overall survival, it did not enhance the strength of efficacy evidence in support of the efficacy claim, from a statistical perspecitve.

Additional analysis using Cox PHR adjusting some sets of covariates could lower p-values for the comparisons. However, the results could be biased mainly due to the exclusion of randomized patients and should be used just as supportive. Therefore, 0.01 is likely the p-value for comparing the two arms in overall survival.

As shown in this review, overall survival was not defined as the primary or secodary endpoint in the protocols. The primary statistical analysis for comparing the two arms in overall survival was not pre-specified. Because of these un-prespecified nature, it is impossible to precisely estimate the operational type I error rate for approving an ineffective product. When a study failed to meet its primary endpoint(s), there was no alpha left for any of the other endpoint analyses. So literally (or from pure statistical point of view), the difference in other endpoints should not be considered statistically significant. This is why it is difficult to interpret the hypothesis test result for overall survival in Study D9901 (p=0.01).

However, with the adjustment for multiple endpoints and multiple analyses, someone may still want to judge that the type I error is controlled under the level of 0.05 or close to this level. This implies that the type I error rate for approving this product may be close to the level of 1/40. However, even if the statistically significant judgment could be considered reasonable, it is certain that the level of making a false claim of effectiveness for this product is still substantially higher than the one in a conventional setting (1/1600).

The key efficacy evidence (difference between the two arms in overall survival) for this BLA is based on post-hoc analyses and the efficacy evidence is not substantial from a statistical perspective.

SUMMARY AND CONCLUSIONS

Study Summary

BLA 125197.0 is an original submission on Provenge (Sipuleucel-T, APC8015) for the treatment of men with asymptomatic metastatic androgen independent prostate cancer. Data from two randomized, double blind, placebo-controlled studies were submitted as the main efficacy evidence under this BLA to support the licensing application.

Both studies failed to meet their primary endpoints and other pre-specified endpoints. However, In Study D9901 with 127 patients, the median overall survival in subjects treated with Provenge was 25.9 months, compared to 21.4 months among placebo subjects. The log-rank test on the difference between the two arms in overall survival resulted in a p-value of 0.01. Study D9902A, an identically designed study with 98 patients, only showed a trend toward improvement in overall survival. It should be noted that overall survival as an endpoint was not defined in both study protocols and the primary statistical analysis for comparing the two arms in overall survival was not pre-specified.

Conclusions

- Based on the results of the statistical analyses of the efficacy data presented by the sponsor and the results of analysis performed by this reviewer, it appears that the two studies provide preliminary efficacy evidence for the treatment.

- However, the key efficacy evidence (difference between the two arms in overall survival) is based on post-hoc analyses. It is difficult to interpret hypothesis test results for overall survival due to the nature of the analyses. The efficacy evidence is not substantial from a statistical perspective.

- The efficacy evidence is not strong enough in support of using Provenge for the treatment of men with asymptomatic metastatic androgen independent prostate cancer.

REFERENCES

1. US FDA. Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drugs and Biological Products. 1998. http://www.fda.gov/opacom/morechoices/industry/guidedc.htm

2. Senn S. Statistical Issues in Drug Development, Section 12.2.7: The two-trials rule. Chichester, England, John Wiley & Sons, 1997: 165-167.

3. Fisher LD. One large, well-designed, multicenter study as an alternative to the usual FDA paradigm. Drug information Journal. 1999; 33: 265-271.